सत्यमेव जयते

INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI

I.A.R I.6.
GIP NLK—H-3 I.A.R.I. --10-5-55—15,000

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 38: 1943
NUMBERS 221–224

Linlithgow Library
Imperial Agricultural Research Institute,
New Delhi

Published Quarterly by the
AMERICAN STATISTICAL ASSOCIATION
WASHINGTON, D. C.
1943

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

CONTENTS OF VOLUME 38

Articles 1, 143, 271, 387
Book Reviews 125, 250, 367, 473
Publications Received 140, 269, 384, 493
Memorial Notes 105, 107, 110, 112, 247, 365, 471
Reports and Official Notices: Report of the Board of Directors, 113; Report of the Secretary, 116; Report of the Committee on Fellows, 117; Reports of the Treasurer and the Auditors, 118–122; List of Committees and Representatives for 1942, 123–124; Committee on Nominations, 366.

ARTICLES, NOTES AND DISCUSSIONS

American Statistical Association, A Special Committee of the. An Appraisal of the U. S. Bureau of Labor Statistics Cost of Living Index 387
Anonymous. Norman John Silberling, 1892–1942 112
Anonymous. The Study of Cycles 248
Berkson, Joseph. Experience with Tests of Significance: Reply to Professor R. A. Fisher 242
Brooks, Edith, see Karsten, Karl
Bruner, Nancy, see Davies, G. R.
Cave, Roy C. Variations in Expenditures where Families of Wage Earners and Clerical Workers are Classified by Economic Level. 445
Cochran, W. G. Analysis of Variance for Percentages Based on Unequal Numbers 287
Comstock, R. E. Overestimation of Mean Squares by the Method of Expected Numbers 335
Cottam, Howard R. Housing Scales for Rural Pennsylvania 406
Cowden, Dudley J. Correlation Concepts and the Doolittle Method. . 327
Daly, Patricia, and Douglas, Paul H. The Production Function for Canadian Manufactures 178
Davies, G. R., and Bruner, Nancy. A Second Moment Correction for Grouping 63
Davis, Shelby Cullom. Coordinating Production for War. 417
Deibler, F. S. Horace Secrist, 1881–1943. 365
Deming, W. Edwards, see Hansen, Morris H.
Deming, W. Edwards, see Tepping, Benjamin J.

Douglas, Paul H., see Daly, Patricia
Durand, E. Dana. A. Manuel Fox, 1889–1942 105
Durand, John D., see Mautz, William H.
Eckler, A. Ross, and Staudt, E. P. Marketing and Sampling Uses of Population and Housing Data 87
Ezekiel, Mordecai. Choice of the Dependent Variable in Regression Analysis—Comments 214
Fisher, R. A. Note on Dr. Berkson's Criticism of Tests of Significance . . 103
Gage, Robert. Contents of Tippett's "Random Sampling Numbers". . 223
Hansen, Morris H., and Deming, W. Edwards. On Some Census Aids to Sampling 353
Hinrichs, A. F. Charles P. Neill, 1865–1942 110
Hirschman, Albert O. On Measures of Dispersion for a Finite Distribution 346
Holmes, Irvin. Some Sampling Uses of Data from the Census of Agriculture 78
Hurwitz, William N., see Tepping, Benjamin J.
Jones, Howard L. Fitting Polynomial Trends to Seasonal Data by the Method of Least Squares 453
Karsten, Karl, and Brooks, Edith. "Retro" Charts 302
Kelley, Truman L. The Evidence for Periodicity in Short Time Series. . 319
Knudsen, Lila F. A Method for Determining the Significance of a Shortage 466
Lotka, Alfred J. Some Reflections—Statistical and Other—on a Non-Material Universe 1
McCormick, Thomas C. An Approach to the Measurement of Farm Population Pressure in Wisconsin 165
Mautz, William H., and Durand, John D. Population and War Labor Supply 31
Moore, Geoffrey H., and Wallis, W. Allen. Time Series Significance Tests Based on Signs of Differences 153
Myers, Robert J. Comparison of Demographic Rates Assumed by the National Resources Committee with Actual Experience 201
Myers, Robert J., and Ober, Harry. Statistics for Wage Stabilization . 425
Northrup, M. Starr, see Webb, John N.
Ober, Harry, see Myers, Robert J.
Palmer, Gladys L. Factors in the Variability of Response in Enumerative Studies 143
Payne, Stanley L., see Webb, John N.
Pierce, Joseph A. Correction Formulas for Moments of a Grouped-Distribution of Discrete Variates 57
Platt, John R. A Mechanical Determination of Correlation Coefficients and Standard Deviations 311
Politz, Alfred. Family Versus Individual in Measurement of Audiences . 233
Rice, William B. Quality Control Applied to Business Administration. . 228
Rogers, Beatrice A., and Tucker, Donald S. Davis Rich Dewey, 1858–1942 107
Shapiro, Solomon. The Distribution of Deposits and Currency in the United States, 1929–1939 438
Shyrock, Henry S., Jr. Internal Migration and the War 16

SIEGEL, IRVING H. Note on a Common Statistical Inequality 217

SIMMONS, C. D. Edward Lewis Dodd, 1875–1943. 247

SMITH, HERBERT G. Utility of Statistical Method in Aerodynamics . . 341

STAUDT, E. P., see Eckler, A. Ross

TEPER, LAZARE. Observations on the Cost of Living Index of the Bureau of Labor Statistics. 271

THORNDIKE, E. L., and WOODYARD, ELLA. The Cost of Living in Cities in Relation to Their Size and Latitude. 238

THORP, WILLARD L. Paul T. Cherington, 1876–1943. 471

TUCKER, DONALD S. The Interest Rate and Saving 101

TUCKER, DONALD S., see Rogers, Beatrice A.

TUCKER, RUFUS S. The Composition of Income and Ownership of Capital by Income Classes in the United States in 1936 187

TEPPING, BENJAMIN J., HURWITZ, WILLIAM N., and DEMING, W. EDWARDS. On the Efficiency of Deep Stratification in Block Sampling. 93

WALLIS, W. ALLEN, see Moore, Geoffrey H.

WAUGH, FREDERICK V. Choice of the Dependent Variable in Regression Analysis 210

WEBB, JOHN N., NORTHRUP, M. STARR, and PAYNE, STANLEY L. Practical Applications of Theoretical Sampling Methods. 69

WHITE, R. CLYDE. The Social Insurance Movement. 358

WICKENS, ARYNESS JOY, and WILLIAMS, FAITH M. Notes on Mr. Teper's Observations 285

WILLIAMS, FAITH M., see Wickens, Aryness Joy

WOODYARD, ELLA, see Thorndike, E. L.

WORKING, HOLBROOK. Statistical Laws of Family Expenditure . . . 43

BOOK REVIEWS

ALTMAN, GEORGE T. *Federal Tax Course.* James D. Paris 373

BACKMAN, JULES, see Stein, Emanuel

BARGER, HAROLD. *Outlay and Income in the United States, 1921–1938.* Everett E. Hagen 368

BARGER, HAROLD, and LANDSBERG, HANS H. *American Agriculture, 1899–1939.* Mordecai Ezekiel 484

BEEBE, GILBERT WHEELER. *Contraception and Fertility in the Southern Appalachians.* Margaret Jarman Hagood 252

BENNETT, KENNETH R., see Pearson, Frank A.

BERG, EVELYN VAN EMDEN, see Brewington, Ann

BERRY, THOMAS SENIOR. *Western Prices before 1861, A Study of the Cincinnati Market.* Ernest W. Williams, Jr. 485

BOWLEY, A. L. (Editor). *Studies in the National Income, 1924–1938.* Milton Gilbert 473

BREWINGTON, ANN, and BERG, EVELYN VAN EMDEN. *The Women Graduates of a Collegiate School of Business.* David Yentis 130

BROOKINGS INSTITUTION. *Refugee Settlement in the Dominican Republic.* Erich W. Zimmermann 254

BUTTERBAUGH, GRANT I. *The Measurement of Business Activity in the Puget Sound Area.* Bervard Nichols 483

CHASE, STUART. *The Road We Are Traveling.* Forrest H. Kirkpatrick . . 377

Clark, Carrie Patton, see Clark, Fred E.
Clark, Fred E., and Clark, Carrie Patton. *Principles of Marketing.* Howard T. Hovde 376
Commodity Statistics, The 1942 Commodity Year Book. Ocsar L. Endler . . 260
Commodity Year Book. Oscar L. Endler 260
Crum, William L., Fennelly, John F., and Seltzer, Lawrence H. *Fiscal Planning for Total War.* Alfred G. Buehler 479
Culliton, James W. *Make or Buy.* Charles E. Young 261
Davis, Hiram S. *Wool and the War.* R. R. Renne 262
Eidenössisches Statistisches Amt. *Neue Schweizerische Sterbetafeln.* Georg Wolff. 125
Encyclopaedia Britannica World Atlas. M. J. Proudfoot 135
Fabricant, Solomon. *Employment in Manufacturing, 1899–1939.* Daniel Creamer 481
Faville, David E., Jones, D. M., and Sonne, R. B. *Merchandise Availability in Utah.* Dilworth Walker 138
Fennelly, John F., see Crum, William L.
Friend, Irwin, see Goldsmith, Raymond W.
Furness, J. W., see Leith, C. K.
Goldsmith, Raymond W., Parmelee, R. C., Friend, I., Gorham, J., and Granby, H. *Distribution of Ownership in the 200 Largest Nonfinancial Corporations.* Ralph C. Epstein and John H. Myers 136
Gorham, James, see Goldsmith, Raymond W.
Granby, Helene, see Goldsmith, Raymond W.
Granby, Helene. *Survey of Shareholdings in 1,710 Corporations with Securities Listed on a National Securities Exchange.* Ralph C. Epstein and John H. Myers 136
Gras, N. S. B. *Harvard Cooperative Society, Past and Present.* Colston Warne 379
Guilford, J. P. *Fundamental Statistics in Psychology and Education.* Martin A. Brumbaugh 266
Haberler, Gottfried. *Consumer Instalment Credit and Economic Fluctuations.* James F. Cusick. 256
Hall, Marguerite F. *Public Health Statistics.* Clyde V. Kiser 253
Hartkemeier, Harry Pelle. *An Introduction to Managerial Business Statistics.* Donald R. G. Cowan 382
Heaney, Norman S. *Public Trusteeship.* C. Herman Pritchett. . . . 487
Holmes, Maurice C. *An Outline of Probability and Its Uses.* E. C. Molina . 267
Horton, Donald C., Larsen, Harald C., and Wall, Norman J. *Farm-Mortgage Credit Facilities in the United States.* John B. Schneider . . . 257
Jacoby, Neil H., see Saulnier, Raymond J.
Jones, Dix M., see Faville, David E.
Katona, George. *War Without Inflation.* A. Smithies 367
Kiser, Clyde V. *Group Differences in Urban Fertility.* Frank Lorimer . . 250
Landsberg, Hans H., see Barger, Harold
Larsen, Harald C., see Horton, Donald C.
League of Nations. *Wartime Rationing and Consumption.* M. K. Bennett. 264
Leith, C. K., Furness, J. W., and Lewis, Cleona. *World Minerals and World Peace.* Wilbert G. Fritz. 375

LEWIS, CLEONA, see Leith, C. K.

MACHLUP, FRITZ. *International Trade and the National Income Multiplier.* Paul A. Samuelson 369

MAGILL, ROSWELL. *The Impact of Federal Taxes.* Richard A. Musgrave . 476

MARGET, ARTHUR W. *The Theory of Prices.* Melchior Palyi. 126

MIGONE, RAUL C. (Editor). *Inter-American Statistical Yearbook, 1942.* Myron E. Andrews. 374

MOLINA, E. C. *Poisson's Exponential Binomial Limit.* J. Neyman . . . 134

NATIONAL BUREAU OF ECONOMIC RESEARCH, CONFERENCE ON RESEARCH IN INCOME AND WEALTH. *Income Size Distributions in the United States, Part I.* Daniel Creamer 475

NIENSTAEDT, L. R. *Economic Equilibrium, Employment and Natural Resources.* Elmer C. Bratt 371

PARMELEE, REXFORD C., see Goldsmith, Raymond W.

PAYNE, WILSON F. *Business Behavior 1919–1922.* V. P. Timoshenko. . . 258

PEARSON, FRANK A., and BENNETT, KENNETH R. *Statistical Methods Applied to Agricultural Economics.* E. L. Butz 265

PETERS, J. (compiled by). *Seven-Place Values of Trigonometric Functions.* Kenneth N. Ogle 268

ROBISON, SOPHIA M. (Editor). *Jewish Population Studies.* Georg Wolff . . 381

ROSTAS, L., see Shirras, G. Findlay

SAULNIER, RAYMOND J., and JACOBY, NEIL H. *Accounts Receivable Financing.* Robert D. Ayars 480

SELTZER, LAWRENCE H., see Crum, William L.

SHIRRAS, G. FINDLAY, and ROSTAS, L. *The Burden of British Taxation.* Tipton R. Snavely 478

SIMON, HERBERT A., et al. *Determining Work Loads for Professional Staff in a Public Welfare Agency.* Constance Hastings 132

SONNE, RICHARD B., see Faville, David E.

STEIN, EMANUEL, and BACKMAN, JULES (Editors). *War Economics.* Alfred G. Beuhler 129

STIGLER, GEORGE J. *The Theory of Competitive Price.* J. P. Miller. . . . 128

STURMTHAL, ADOLF. *The Tragedy of European Labor, 1918–1939.* Robert I. Crane 488

TAX INSTITUTE. *Tax Yields: 1941.* Henry S. Bloch 477

*TEMPORARY NATIONAL ECONOMIC COMMITTEE

TOMASEVICH, JOZO. *International Agreements on Conservation of Marine Resources.* Reginald H. Fiedler 486

TRELOAR, ALAN E. *Random Sampling Distributions.* Margaret Merrell . . 133

TWENTIETH CENTURY FUND. *How Collective Bargaining Works.* Dale Yoder 131

VON MERING, OTTO. *The Shifting and Incidence of Taxation.* Carl S. Shoup . 372

WALKER, HELEN M. *Elementary Statistical Methods.* Laverne K. Burke . 489

WALL, NORMAN J., see Horton, Donald C.

WILKS, S. S. *Mathematical Statistics.* Abraham Wald 491

ZAPOLEON, MARGUERITE W. *Community Occupational Surveys.* Harold F. Clark. 490

---

* For TNEC Monographs Nos. 29 and 30, see Goldsmith and Granby.

BOOK REVIEWERS

Andrews, Myron E., 374
Ayars, Robert D., 480
Bennett, M. K., 264
Bloch, Henry S., 477
Bratt, Elmer C., 371
Brumbaugh, Martin A., 266
Buehler, Alfred G., 129, 479
Burke, Laverne K., 489
Butz, E. L., 265
Clark, Harold F., 490
Cowan, Donald R. G., 382
Crane, Robert I., 488
Creamer, Daniel, 475, 481
Cusick, James F., 256
Endler, Oscar L., 260
Epstein, Ralph C., 136
Ezekiel, Mordecai, 484
Fiedler, Reginald H., 486
Fritz, Wilbert G., 375
Gilbert, Milton, 473
Hagen, Everett E., 368
Hagood, Margaret J., 252
Hastings, Constance, 132
Hovde, Howard T., 376
Kirkpatrick, Forrest H., 377
Kiser, Clyde V., 253
Lorimer, Frank, 250
Merrell, Margaret, 133
Miller, J. P., 128
Molina, E. C., 267
Musgrave, Richard A., 476
Myers, John H., 136
Neyman, J., 134
Nichols, Bervard, 483
Ogle, Kenneth N., 268
Palyi, Melchior, 126
Paris, James D., 373
Pritchett, C. Herman, 487
Proudfoot, M. J., 135
Renne, R. R., 262
Samuelson, Paul A., 369
Schneider, John B., 257
Shoup, Carl S., 372
Smithies, A., 367
Snavely, Tipton R., 478
Timoshenko, V. P., 258
Wald, Abraham, 491
Walker, Dilworth, 138
Warne, Colston, 379
Williams, Ernest W., Jr., 485
Wolff, Georg, 125, 381
Yentis, David, 130
Yoder, Dale, 131
Young, Charles E., 261
Zimmermann, Erich W., 254

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

MARCH, 1943 · $1.50 PER COPY · $6.00 PER ANNUM · VOL. 38 · NO. 221

CONTENTS

ARTICLES

Some Reflections—Statistical and Other—on a Non-Material Universe . . . ALFRED J. LOTKA 1
Internal Migration and the War . . . HENRY S. SHRYOCK, JR. 16
Population and War Labor Supply . . . WILLIAM H. MAUTZ AND JOHN D. DURAND 31
Statistical Laws of Family Expenditure . . . HOLBROOK WORKING 43
Correction Formulas for Moments of a Grouped-Distribution of Discrete Variates . . . JOSEPH A. PIERCE 57
A Second Moment Correction for Grouping . . . G. R. DAVIES AND NANCY BRUNER 63
Practical Applications of Theoretical Sampling Methods . . . JOHN N. WEBB, M. STARR NORTHROP, AND STANLEY L. PAYNE 69
Some Sampling Uses of Data from the Census of Agriculture . . . IRVIN HOLMES 78
Marketing and Sampling Uses of Population and Housing Data . . . A. ROSS ECKLER AND E. P. STAUDT 87
On the Efficiency of Deep Stratification in Block Sampling . . . BENJAMIN J. TEPPING, WM. N. HURWITZ, AND W. EDWARDS DEMING 93
The Interest Rate and Saving . . . DONALD S. TUCKER 101
Note on Dr. Berkson's Criticism of Tests of Significance R. A. FISHER 103
A. Manuel Fox, 1889–1942 . . . E. DANA DURAND 105
Davis Rich Dewey, 1858–1942 . . . BEATRICE A. ROGERS AND DONALD S. TUCKER 107
Charles P. Neill, 1865–1942 . . . A. F. HINRICHS 110
Norman John Silberling, 1892–1942 . . . 112

REPORTS OF THE ASSOCIATION

Report of the Board of Directors . . . 113
Report of the Secretary . . . 116
Report of the Committee on Fellows . . . 117
Report of the Treasurer . . . 118
Report of the Auditors . . . 119
List of Committees and Representatives for 1942 . . . 123

Continued on next page

BOOK REVIEWS

Backman, Jules, see Stein, Emanuel

Berg, Evelyn Van Emden, see Brewington, Ann

Brewington, Ann, and Berg, Evelyn Van Emden. *The Women Graduates of a Collegiate School of Business.* David Yentis 130

Eidenössisches Statistisches Amt. *Neue Schweizerische Sterbetafeln.* Georg Wolff 125

Encyclopaedia Britannica World Atlas. M. J. Proudfoot 135

Faville, David E., Jones, D. M., and Sonne, R. B. *Merchandise Availability in Utah.* Dilworth Walker 138

Friend, Irwin, see Goldsmith, Raymond W.

Goldsmith, Raymond W., Parmelee, R. C., Friend, I., Gorham, J., and Granby, H. *Distribution of Ownership in the 200 Largest Nonfinancial Corporations.* Ralph C. Epstein and John H. Myers 136

Gorham, James, see Goldsmith, Raymond W.

Granby, Helene, see Goldsmith, Raymond W.

Granby, Helene. *Survey of Shareholdings in 1,710 Corporations with Securities Listed on a National Securities Exchange.* Ralph C. Epstein and John H. Myers 136

Jones, Dix M., see Faville, David E.

Marget, Arthur W. *The Theory of Prices.* Melchior Palyi 126

Molina, E. C. *Poisson's Exponential Binomial Limit.* J. Neyman . . 134

Parmelee, Rexford C., see Goldsmith, Raymond W.

Simon, Herbert A., *et al. Determining Work Loads for Professional Staff in a Public Welfare Agency.* Constance Hastings 132

Sonne, Richard B., see Faville, David E.

Stein, Emanuel, and Backman, Jules (Editors). *War Economics.* Alfred G. Buehler 129

Stigler, George J. *The Theory of Competitive Price.* J. P. Miller . . . 128

*Temporary National Economic Committee

Treloar, Alan E. *Random Sampling Distributions.* Margaret Merrell . 133

Twentieth Century Fund. *How Collective Bargaining Works.* Dale Yoder 131

Publications Received 140

---

* For TNEC Monographs Nos. 29 and 30 see Goldsmith and Granby.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 221 MARCH, 1943 *Volume 38*

SOME REFLECTIONS—STATISTICAL AND OTHER—ON A NON-MATERIAL UNIVERSE[*]

BY ALFRED J. LOTKA
Metropolitan Life Insurance Company

As YOU KNOW, the exigencies of the war caused us to abandon our plans for the regular annual convention in Cleveland.

It had been my expectation to address you informally at a luncheon meeting, in that genial mood which comes as a sort of blessing after the meal—when the critical faculties are just a little dulled; when minor blemishes pass unobserved and major defects are more easily forgiven.

But now I must transcribe my message. The lighter cargo of the spoken word must be brought to you in the heavier vehicle of the printed page. Nevertheless, light cargo it shall remain, for I have it on good authority that people can be fatigued without knowing it, but they cannot be bored without knowing it. And I shall try not to bore you.

The statistician is in one respect—and in only one respect, I fear—a fortunate person. The Universe is his topic. He can discourse about anything and everything. I recall a drawing in an old American magazine, since discontinued. It pictured a class of undergraduates sitting for an examination. The professor was writing the questions on the blackboard. The first question was: "What is, and why is it?" I think this must have been a class in statistics.

It is not my purpose to attempt here a complete answer to the professor's question. I have in mind to cover only one half of the Universe, a certain special half. Classification is an art that allows considerable freedom to those who practice it. Taking advantage of this freedom I shall for the present occasion divide all statistics into statistics of material, and statistics of non-material things; the latter, as we shall all agree, being by no means immaterial things. It is about these non-material statistics that I ask you to think with me for a little while.

---

[*] This paper was to have been given as the Presidential Address at the 104th Annual Meeting of the American Statistical Association, to have been held at Cleveland, Ohio, December 1942.

In response to a request from the Office of Defense Transportation to avoid adding to the traffic congestion during the holiday season the meetings were cancelled. This makes the first gap in an otherwise unbroken series of 103 Annual Meetings of the Association.

16637

Although the non-material Universe is ever-present with us, or perhaps because of this very fact, we are inclined at times to ignore or even to question its presence. It is true that much of it is trivial. I refer to the Universe of thought. Every one of the millions of normal adult human inhabitants of this globe, day in and day out, throughout his waking hours, spins a thread of thought. The thought just now in Mrs. Jones' mind may be nothing more important than the question where Mrs. Smith bought her hat; while Mr. Smith may be thinking of the $10 he won last night at cards from Mr. Jones. Statistics, even of such trivial thoughts, may not be wholly without value. Psychologists and psychiatrists may have an interest in them. But not all thought is trivial. There is much thought about the intensely practical affairs of life. And there is an aristocracy of thought. Even this is not immune from attack by the statistician.

Anyone may, if he so chooses, conduct a statistical inquiry into his own thoughts, dealing directly with the non-material data themselves. But if the inquiry is to be made more objective, to cover thoughts of many individuals, some form of materialization of these thoughts must intervene; the statistics then are no longer quite strictly statistics of non-material things, but actually statistics of the material symbols that stand for these things. Such symbols may be spoken or written words or sentences, or they may be corresponding musical performance or notation.

They may be spontaneous expressions, or they may be solicited answers to questionnaires or polls. As to these latter, it is interesting to note that here we meet with a phenomenon long familiar in this field, but only lately recognized in the domain of physical observations also, namely that the very act of interrogating alters the state of the object addressed, so that strictly speaking we can never by questioning be really sure of discovering the exact and true answer to the question posed. We note with regret that illustrations of this fact are easily gathered in any court of justice.

The statistical analysis of expressed thought has a certain interest and utility from various angles. The practical man, who is primarily concerned with making life possible and economically profitable, has had occasion to play with statistics of both the spoken and written language. The printer, unless he casts his type as he goes, wants to know in what proportion to stock the various letters of the alphabet. For quite different reasons the cryptanalyst, in solving intercepted cipher messages, is interested in the frequency of these letters in different languages.

The telephone engineer has explored[1] the frequency of the principal words in colloquial English, since he is vitally interested in securing the best possible transmission, and this obviously requires a knowledge of the composition of spoken sounds.

While his investigations were made primarily for purposes of practical phonetics, in its application to telephone transmission, incidentally some of the by-products throw interesting sidelights in other directions. Our language is a population of words in which there is a sort of social stratification of function. Nouns, names of concrete things or abstract concepts, were found to be by far the most varied, nearly one half of the *different* words used—to be exact, 1,029 out of 2,240, in 500 telephone conversations. The number of *different* verbs were less than half this, namely 456. This, of course, is not accident, but has a meaning. The variety of *things* appearing in the universe of conversation analyzed was greater than the variety of *happenings* in which these things were involved. This is not surprising, for a typical statement or clause contains, associated with its verbal predicate, a subject, an object (if the verb is transitive), and in many cases further nouns qualifying the subject or object or the verb itself by an adverbial phrase. However, in these telephone conversations, the *total* occurrences of verbs (including repetitions, but exclusive of auxiliary verbs), namely 12,550, actually exceeded slightly the *total* occurrences of nouns, 11,660. While the same verb was repeated on an average eleven times in the total of nearly 80,000 words, the same noun was repeated on an average less than three times. The relatively low proportion of total nouns is probably peculiar to telephone conversations, which on the other hand, are for obvious reasons very rich in pronouns, the words "I" and "you" heading the list of all words.[2]

All this suggests that certain old-fashioned school exercises in grammar can be developed into something more interesting and of greater significance, in the form of *quantitative parsing*. This sort of study practically amounts to a statistical appraisal of the furniture of the mind, insofar as this is reflected in language.

True, not all words necessarily reflect useful mental furniture. So it has been remarked that the *ether* of the physicist of the 19th century may have been nothing more than the subject for the verb to undulate.

---

[1] N. R. French, C. W. Carter, Jr., W. Koenig, Jr., "The Words and Sounds of Telephone Conversations," *Bell Telephone System Technical Publications*, June 1930; N. R. French and W. Koenig, Jr., *Journal Acoustical Society of America*, 1929, p. 110; H. Fletcher (Bell Telephone Laboratories), *Speech and Hearing*, Van Nostrand, 1929.

[2] Compare also Vernon Lee, *The Handling of Words and Other Studies in Literary Psychology*, 1923; see especially Chapter VI, pp. 187 et seq., L. W. Faucett, *A Study of English Word Values*.

The words mind, soul, may be just grammatical conveniences, subjects for such verbs as to think, to feel, etc. Such words, if not very useful, may nevertheless be relatively harmless. But they may also cause confusion, or worse, a sort of pseudo-scientific ponderosity. There is such a thing as "attempting to conceal truism behind pretentious terminology."

Singularly enough, the trap of verbiage, from which writers aspiring to profundity have not always escaped, has been quite clearly perceived by the comedian and the wit. So Molière in his *Le malade imaginaire*, has the candidate for a Doctor's degree in medicine interrogated

"Quare opium facit dormire?"

to which he gives the illuminating reply

"Quia est in eo virtus dormitiva."

More lately, and in less classic vein, that versatile author *Anon* has defined the elusive term "life"—whose definition great biologists have discreetly sidestepped:

There was a young man of Cadiz
Who affirmed Life is what it is;
For he early had learnt,
If it were what it weren't
Life would not be what it is.

The fact is that our language comprises, among others, two great classes of words. The first of these is what might be called *Berlitz words*, which can be defined, as in the Berlitz method of teaching foreign languages, by pointing to an object. The second class is what I might call *context words*, the meaning of which (if any) can be ascertained only by examining the contexts in which they occur. Some of these latter words, undoubtedly, are mere counterfeits, or even slugs, which have been smuggled into the circulation of linguistic currency. Such, for example, is the word "ghost."

But these words of doubtful legitimacy are after all a minor part of our speech and writing. They represent a chapter in the morbid anatomy of language. Exaggerations regarding these and other linguistic failings, to which a certain overpopularization of semantics has led several authors,[3] will be avoided in any well-balanced study, for which a standard has been set by Helen Eaton's[4] scholarly work. In an appendix to her *Semantic Frequency List for English, French, German and*

---

[3] For criticism of this, see M. Schlauch, "Semantics as Social Evasion," *Science and Society*, 1942, Vol. 6, p. 315. See also, by the same author, *The Gift of Tongues*, Modern Age Books, 1942.

[4] H. Eaton, *Semantic Frequency List for English, French German and Spanish* (first six thousand words), Chicago University Press, 1940.

Spanish, she brings a statistical tabulation of the 6,000 most common words classified in natural categories, which may be thought of as the basis for an analysis of our conceptual universe.

The parallelism between that universe and the universe of language is admittedly incomplete. For one thing, our English language, in particular, is extremely rich in synonyms, two or more words denoting the same or closely similar concepts. A comparative study of this multiplication of terms in different languages would be of interest. We recall, for instance, that the French have no word for *home*, but must get along with a *house* or a *hearth*. The Frenchmen can distinguish only by mien, not by specific words, between his sentiment for Colette and his appreciation of a good cigar, or even of a piece of chocolate. But the English language, too, has its lacunae. The French *vieillard*, the German *Greis*, the Latin *senex*, and the Greek *γέρων* have no counterpart in English; though according to our custom we have appropriated the Greek root in naming a branch of science, gerontology, recently come into prominence on account of the aging of our population to which our vital statisticians have been drawing attention.

Such lacunae in various tongues may be fortuitous. But wealth of synonyms is clearly enough ascribable to definite causes. English, undoubtedly, owes its wealth in this field largely to its polyglot origin. But other factors also enter, irrespective of any particular language. Words with painful associations tend to run to synonyms. So in English we avoid reference to death and speak of passing away. We have three words, graveyard, churchyard, cemetery; the German speaks of Kirchhof, Friedhof, Gottesacker. Examples could easily be multiplied. For more or less disreputable things slang often supplies a whole string of synonyms. Another motive for multiplying synonyms seems to be the desire to achieve and surpass superlatives: big, large, great, huge, immense, enormous, gigantic, monstrous, jumbo, etc.; and at the other extreme, small, little, *petite*, pygmy, puny, minute, diminutive, tiny, wee, etc.

To the creative writer, the *littérateur*, this wealth of synonyms, allowing as it does fine shadings and distinctions of meaning, is a treasure. But there is also a diametrically opposite attitude. To a certain type of mind the ideal language would be one in which each word or symbol would have one single meaning, and in which for each meaning there would be a single word. On the construction of such a language, free from all ambiguity and redundancy, certain mathematicians have labored, with perhaps somewhat uncertain success. And there is also what might be called the economic philologist, who wishes to undo the mischief of the Tower of Babel. Attempts to substitute highly efficient artificial languages for the inefficient natural tongues which, like Topsy, "just grew," do not seem to have met with much success, and in fact have built a little Tower of Babel of their own.[5] Ogden's proposal of *Basic English,* which manages somehow to say most of what is worth saying, by the use of only about 850 words, seems to be a more practical venture. A perhaps minor defect from which it suffers is a strange sort of irreversibility, in that any English-speaking person can read, but by no means write it. To do the latter he must first learn what words *not* to use—a singularly backhanded method of learning a language. The foreigner must learn some 800 English words. But the native must "unlearn" many thousands.

I have remarked that the practical man, primarily concerned with making life possible and economically profitable, has had occasion to concern himself with statistics of language. But the artist and the scholar also, whose chief contribution is to make life pleasant and interesting, have paid some attention to, and certainly have supplied much material for, the statistics of thought expressed in sound and symbol. The statistician has not hesitated to apply his technique to Horace's Odes, to Greek statues or to literary style. Among the latest contribution in the latter class are two articles in *Biometrika,* the first by G. U. Yule,[6] followed soon after by a more detailed article by C. B. Williams.[7] The latter compared the frequency distribution of sentence length in Bernard Shaw, G. K. Chesterton and H. G. Wells. A good many years ago, I myself started what I might call an Inventory of the Poet's Mind. From concordances of Wordsworth, Milton and Keats, I noted the frequency of a considerable number of words. For example, the relative number of mentions of father, mother, sister, brother, in Wordsworth was in the ratio of 37 to 35 to 10 to 18; in Milton the ratio was 60 to 20 to 7 to 13; in Keats, 25 to 16 to 28 to 31. I did not continue this investigation when I found that this kind of thing had already been done by others. So, for example, Niceforo in his *Méthode Statistique,* quotes a study of the frequency of mention of various colors in the sonnets of Baudelaire, and gives many examples of the application of statistics to literature and art, among others a study of the distribution of sentence length in Balzac.

---

[5] Both the First World War and the present war have been the occasion for a revival of the movement for an international language. In this connection see *Nature,* 1942, Vol. 150, pp. 622 and 718; 1943, Vol. 151, p. 80.

[6] G. U. Yule, "On Sentence Length as a Statistical Characteristic of Style in Prose: With Application to Two Cases of Disputed Authorship," *Biometrika,* 1939, p. 363.

[7] C. B. Williams, "A Note on the Statistical Analysis of Sentence Length as a Criterion of Literary Style," *Biometrika,* 1940, p. 356.

A recent contribution is a statistical study of sound patterning in poetry by B. F. Skinner,[8] who has also examined the frequency distribution of response words in association tests. Plotting the logarithms of frequencies as ordinates against the logarithms of ranks as abscissae (the most frequent word has rank one, etc.), he found that over the major part of the range the data fell essentially along the straight line. This relation, however, is not restricted to words collected by the process of association. It appears to be the general characteristic of words in actual use. It had been previously observed and discussed by G. K. Zipf,[9] who has made interesting suggestions regarding the underlying causes of the phenomenon. It is to be noted that relations of this kind are observed in a great variety of seemingly unrelated phenomena including, among others, the frequency distribution of scientific productivity and the relation between rank and population of the cities of the United States and of certain other countries.[10]

The population of words, like any population, grows by new accessions in excess of eliminations. The new accessions, in modern times, have largely been names for new objects created by man's ingenuity. Curiously enough, the same agency has caused the virtual disappearance of words. How many of the younger generation know what is a pair of snuffers? Some ingenious person hit on the plan of so braiding the candlewick that it curled as it burned, its tip projecting into the oxidizing zone of the flame, to be there consumed as it goes, instead of having to be trimmed with a special kind of scissors. And so, while snuffers may be found occasionally in an antique shop or a museum, the phrase "pair of snuffers" is today little more than a dictionary entry. Words can also be lost simply by being found superfluous. The word "shall" is thus in process of becoming extinct in the Anglo-American language.

The ideal dictionary would not merely list words with their meaning, but would give statistical data regarding their relative frequency in spoken and in written language.[11] Successive editions, published at considerable intervals, would give a quantitative record of the history of the language. This would have been of particular interest for the period of rapid transformation of Latin into French and the other Romance languages. In comparison with this and corresponding transitions in other tongues, modern speech is relatively at a standstill, anchored by the weight of the printing press.

---

[8] B. F. Skinner, "A Quantitative Estimate of Certain Types of Sound Patterning in Poetry," *American Journal of Psychology*, 1941, p. 64; S. W. Cook and B. F. Skinner, "Some Factors Influencing the Distribution of Associated Words," *Psychological Record*, 1939, p. 178.

[9] G. K. Zipf, *Selected Studies of the Principle of Relative Frequency in Language*, Harvard Press, 1932; *The Psychobiology of Language*, Houghton Mifflin, 1935, second edition, 1939; *National Unity and Disunity*, Principia Press, 1941, pp. 398–399.

[10] A. J. Lotka, *Science*, 1939, Vol. 94, p. 164; *Journal Washington Academy of Science*, 1926, Vol. 16, p. 317; *Elements of Physical Biology*, 1925, p. 306.

[11] This is not altogether a dream of the future. Helen Eaton, *loc. cit.*, p. xiii, quotes F. W. Kaeding, "Häufigkeitswörterbuch der deutschen Sprache," containing 79,716 words divided into 11 categories.

Studies of thought expressed in speech or writing can remain on the purely descriptive level, or on the lowest rung of statistical technique, classification and enumeration. They may assume a literary aspect, as in the writings of some of our modern or no longer quite modern writers, who have sought to fix on the canvas of literature something of the ill-assorted stream of consciousness.

But thoughts are not independent events, to speak in the language of the statistician; or, in the language of the physiological psychologist, they are conditioned reflexes. In the evolution of human culture, order has been established at least in some realms of thought. The conditioning by which one thought leads to another has become systematized. Fundamental to this scheme is the principle that competent thought must be free from contradictions. At the same time certain realms of thought have acquired the character of an organic whole, with results of the greatest practical importance, but sometimes perhaps with a deceptive appearance of being all-inclusive. Scrutiny reveals the fact that the reason for the coherence of this type of systems of thought is that they involve a circle which is by no means vicious: Contradictions cannot arise in a system which defines a cello as a large violin, and a violin as a small cello—to use an illustration cited, if my memory serves, by the late J. W. N. Sullivan. If the operations in such a scheme lead to inevitably consistent results, and in that sense to truisms, this does not by any means deprive the scheme of practical utility, as every physicist knows. Whether all competent thought can be united in one coherent whole, or in two or more distinct, but each separately coherent parts, remains today an open question. At any rate, it seems that the best minds come nearest to achieving such coherence; in such minds, as has been pointed out by Professor C. J. Keyser,[12] thoughts seem to approach the structure which the mathematician speaks of as the "group property."

As to the laws of thought which bring about this pleasing harmony, for a long time they had been supposed to have absolute precision, and so to be outside the realm of statistics. More recently the sanctity of these supposedly absolute laws has been questioned. Systems have been proposed in which the sharp lines of classical logic give way to the blurred margins of probabilities. As yet the merits of these tentatives are a matter of dispute. This is a battlefield in which I find myself a non-combatant in no-man's land, where any overt action on my part would expose me to crossfire fore and aft.

---

[12] C. J. Keyser, *Mathematical Philosophy*, Dutton, 1922, p. 227.

Recorded thought is preponderantly thought of the past. But there is other thought—often unexpressed, and for that very reason of particular importance: thought of the future, *purposive* thought. In this the statistician may well be said to have a special interest. For the ultimate use of all statistics, except the "Believe it or not" type, is prediction.

Now prediction is of two kinds. On the basis of past observations the astronomer predicts eclipses of the sun and the moon, the position of the planets, and so forth. He is a sort of elongated dachshund, whose nose sniffs tomorrow, while his tail wags yesterday.

But there is another, totally different kind of prediction. Each of us has his own private field of prediction. I can predict events that no one else can predict. I can predict that when I have completed this sentence I shall get up from my chair and walk around the room. No one else could have predicted that. This private field of prediction consists of the events which we purpose, the events which we successfully will. The English language has embedded this fact in its very structure, by appropriating the vocable 'will" as the auxiliary verb to characterize tense. But note the significant fact that this auxiliary has a companion of somewhat different complexion: the word "shall," which in some contexts implies a conflict of wills, the imposition of one person's will upon the future actions of another.

Now, in this particular part of the universe of thought also—*purposive* thought—there is much that is trivial, much that is chaotic: Purposes and cross-purposes, will contending with will to null effect, where harmony might have produced important collective resultants. For here also organization is possible. Some kind of order can be brought into the assortment of purposes that motivate human actions. Such ordering influences have indeed been exerted over a wide field, ranging from the sublime teaching of great ethical leaders, through the relatively innocuous drivel of radio sales talk, down to the diabolical propaganda of a herd of gangsters, and their indoctrination of plastic youth with calculated deviltry. A certain, perhaps mistaken delicacy, has allowed little play to statistical investigation of the degree of success of ethical teaching. Possibly a healthy, although humiliating lesson might be learned from such an investigation. No delicacy stands in the way of investigating the effectiveness of advertising as influencing consumer demand, and studies along that line are today a common-place—though one wonders how so much piffling advertising has survived the scrutiny.

As for propaganda, though experimental studies have been made under "laboratory conditions," propaganda in the wild state of nature, and in its most pernicious form, is inherently difficult of access to statistical examination, for, to quote C. H. Hoffer,[13] "Propaganda is the process of influencing attitudes in a veiled manner, regarding matters of controversy within a society. Since it is veiled and partisan in nature, propaganda retards the formation of intelligent public opinion, and thus is a menace to democracy." The very obliquity and secretive character of propaganda tend to withdraw it from examination in the honest light of day.

The indoctrination of youth is a somewhat debatable subject, though the horrible examples of Naziism and Fascism and other "isms" certainly give us pause. Some may agree with me and some may not, if I say that the function of schools and colleges, aside from cultivating an appreciation of the humanities, is to teach facts and methods, thought processes, and the faculty of forming independent opinions; but that it is none of the functions of such institutions to inculcate ready-made opinions. Such indoctrination is abhorrent both to the spirit of scientific truth and to all sense of rational and honest citizenship.

But what has this to do with statistics? The answer is obvious: The whole mechanism of democratic government is statistical. The election of our governing bodies is a statistical experiment on a national scale. As to its results, the experiment is not always an unqualified success. We have not reached perfection. A variety of unofficial polls which have come into vogue perhaps point a direction in which lies improvement. Certainly, as a gauge of popular sentiment of the nation's thoughts and purposes, these polls are sometimes more representative than the actions of the elected representatives of the people. That such polls are by no means infallible, and indeed are not wholly free from danger of abuse, must be admitted. But abuse does not disqualify competent use. The unofficial inventories of the nation's mind may contain the germ of important developments.

I have remarked that statistics of non-material things are by no means immaterial. This may be one of those propositions that need only to be stated to be at once accepted as true. Yet I wonder whether it is generally realized how greatly, and on what scale, the non-material universe of thought, and especially of purposive thought, has affected the course of material, physical happenings. Actually it has been nothing less than a geological agency, a phenomenon of planetary dimensions. This is a topic that could be expanded at length, but here it must suffice to give one or two illustrations. It has been estimated[14] that the amount of carbon dioxide annually thrown into the air by the combustion of coal by man is about the same as the total carbon dioxide produced annually in the respiration of the earth's total animal population, and is of the order of one three-hundredth of the total carbon dioxide in the air. It represents a quantity of carbon which surely must have taken hundreds of thousands of years to accumulate as coal in the earth's crust.

---

[13] C. H. Hoffer, "A Sociological Analysis of Propaganda," *Social Forces*, 1942, p. 445.

The case of nitrogen is even more remarkable. For many years the increasing agricultural needs of the rapidly growing human population had largely been met by consuming supplies of nitrates accumulated through the ages in the saltpeter beds of Chile. As the possible exhaustion of this supply began to come into view, less than 70 years after it had been opened to exploitation (1831), men turned their thoughts on this problem, and human ingenuity unlocked the practically inexhaustible supply of atmospheric nitrogen. By 1935 about 75 per cent of the total world's production of bound nitrogen was obtained from this source.[15]

But it is not only, and perhaps not primarily, for the purposes of the peaceful conduct of life, that human ingenuity had been applied. Certainly a powerful stimulus to the early development of the nitrogen fixation industry was the desire in certain quarters to provide for the needs of war, to acquire a supply of raw materials for explosives, independent of the Chilean nitre beds.

It is here that the contrast is seen between the highly effective organization of reflective thought, logic, mathematics, theory of probability and their applications, on the one hand, and, on the other, the woefully incompetent organization of purposive thought. The fact is, that the human species as it stands today is sick. It suffers from a complaint which is more easily diagnosed than cured, because it is congenital.

I say that the human species is sick. It is sick in the sense that any species afflicted with an assortment of propensities that impedes it in the struggle for survival—that perhaps threatens its survival—can certainly be said to be sick. If we want a name for this sickness we might call it pernicious orthogenesis.

Now orthogenesis is one of those big words which we prefer to use sparingly. I have borrowed it from biology, and even there it is not in very good repute. But for my present purpose it is a useful word, though here, addressed to a mixed audience, it may call for a little explanation. Some biologists have thought that the evolution of certain species may have proceeded, not by a mere preferred survival of more advantageous departures from the mean, but by an inherent tendency of certain species, quite aside from external influences, to drift in a certain direction. By a sort of inherent biological proclivity, certain species seem to have been predestined to success, or foredoomed to failure. So, for example, the giant saurians of Jurassic time perhaps became the victims of their fatal propensity for growing bigger and bigger.[16]

---

[14] A. Bethe et al., *Handbuch der normalen und pathologischen Physiologie*, 1927, Vol. 1, pp. 702, 721.

[15] E. R. Riegel, *Industrial Chemistry*, 1937, pp. 111, 112.

Whether such orthogenesis by physiological processes or anatomical features actually has taken place is a matter of dispute among biologists. In this dispute we need not take part, because the existence of the kind of orthogenesis with which we are here concerned will not be disputed.

The fundamental characteristic of orthogenesis is that it proceeds in a set direction, regardless of the utility or the fatality of that direction. For that very reason orthogenesis becomes a powerful factor in evolution, because it can proceed much more rapidly than the gradual accumulation of variations or mutations from generation to generation. Indeed, it can proceed at a rate which, on a geological scale, is nothing short of explosive suddenness.

This is precisely what has happened to the human species in the past few centuries, after a period of preparation extended over perhaps an equal number of millennia.

Curiosity, that ingrained trait of human nature, is the mainspring of this type of orthogenesis. In these days it is rather out of fashion to extoll the virtues of pure science, or even to attempt any distinction between pure and applied science. But it remains a fact of superlative importance that men have questioned Nature from simple curiosity, and have wormed out many of her secrets, with no conscious intent to make so-called useful application of the knowledge gained—or indeed without suspecting that it might bring catastrophe crashing down upon the peoples of the earth. These applications, favorable or unfavorable to the species, have come *after* the scientific discovery—and not infrequently a long time after, figuring in the ordinary human time scale. Indeed, some of these applications were not possible until considerable progress had been made in the unprejudiced interrogation of Nature. This is the typical characteristic of orthogenesis—continued movement in a set direction, independently of the utility or harm of the result, until a stage is reached where selection has something tangible to lay hold on, and to operate either for greater survival, or, it may be, for destruction.

---

[16] A. J. Lotka, *Physical Biology*, 1925, pp. 296, 378.

This then is the sickness with which the human species is sick:

By a process of orthogenesis, that is, through the accumulation of knowledge and skills from generation to generation, man has developed marvelous powers for the individual to achieve his purposes. But he has failed, and failed very badly, to organize these purposes for their survival value. Man's evolution, in this respect, has been lopsided, and there is evidence in the rocks that such lopsided development is a danger signal for the species.

In these times of grave historical events, even an informal address from your retiring president cannot fail to be projected against the background of the weighty issues that are occupying the minds of all responsible men and women. The winning of the war is necessarily the task of our military leaders and the forces bravely carrying out their strategies. The winning of the peace is the concern of us all, but more especially of those who will give orientation to purposive thought.

It is an unfortunate fact that to guide men into competent purposive courses is far more difficult than to teach them the use of ingenious contrivances for attaining such ends as they may set themselves. We give a man an automobile and show him how to turn a switch, shift a lever, step on the gas, and away he goes. He needs no persuasion that the scheme works—the demonstration would convince even that legendary skeptic who, on seeing a giraffe, remarked "there ain't no such animal."

But in guiding men's purposes we have no such compelling demonstrations to present as a means of persuasion. Added to the difficult problem of *devising* efficient policies, there is the still more difficult one of overcoming resistance to their *adoption*. This resistance is twofold. The people must first be persuaded of the need and the reasonable effectiveness of the proposed policies; second, what is even more difficult, they must be induced to sacrifice certain personal advantages in the pursuit of purposes transcending individual interests.

And there is another important and fundamental difference between the influence of organized *reflective* thought in determining group or collective activity of human society and, on the other hand, the role of organized *purposive* thought in directing group behavior: It does not require any very high intelligence for the individual of the rank and file to profit by the accomplishments of our technical experts and scientific geniuses. In this connection, a series of photographs reproduced in the newspapers the other day told an instructive fable. It showed a genial chimpanzee going through the orthodox motions of lighting a cigar and proceeding thoroughly to enjoy a good smoke. But it was a little artistic detail that caught my particular attention: This chimp, we are informed, always begins by carefully stripping the brand label from the cigar. He does not, of course, know why he does this except that he has seen it done. I am afraid that many of our actions have not a very much sounder basis. And so long as we are dealing with what I might call the mechanics of everyday life, this very diluted participation in the management of the world works satisfactorily enough. We can leave the thinking to others, and just push buttons, relying on the mysterious mechanism behind them to work according to schedule.

But this sort of thing is quite inadequate for the direction of collective human activity by and for organized *purposive* thought. In a democracy, at any rate, the individual must personally and understandingly participate in the management of affairs. This makes high demands on the general level of intelligence, wisdom, and enlightened patriotism under our form of government.

And here let us do our thinking as statisticians, and above all as realists. Evil men and groups of such there will presumably always be among the earth's population. No scheme which fails squarely to face this fact can receive our serious consideration. And the danger probably can never be eliminated that, from time to time, an evil genius may give false and wicked orientation to the purposive thought of multitudes.

I say let us do our thinking as statisticians, for the problem with which we have to deal is fundamentally a problem of numbers.

There are certain established, and on the whole very effective, methods of dealing with evil men on a civil scale. There are punishments as deterrents, and there are ways of restraining and constraining the unruly, if need be with capital finality. But fundamental in these methods is a numerical situation: The law-abiding and law-enforcing element, in civil life, in national life, immensely outnumbers the miscreants.

Unfortunately, when we seek to translate into international procedures the principles and methods of civil law enforcement, we are met at the very start with a complete lack of parallelism. Nations, unlike individuals, are not counted in millions, not even in hundreds; and those of dominating power, not even in tens. This completely upsets the parallel.

When an issue arises between nations, it is not a case of a handful of men against millions, but at best of one nation against a very limited number of others—four or five perhaps if the contention is among the world's great powers. What is more, the extreme contrast of right and wrong in the present contest must not obscure to our sight the fact that ordinarily issues are not thus drawn with indisputable clarity, and if sides are taken, say, by five powers or their representatives, as to the merits in a dispute, the split may not be four to one, but three to two, in favor of one or the other contention. Such a split as this will in nowise guarantee a decision without the use of force, or a wise or good decision, with or without the use of force. And if the use of force should be enjoined by some policy previously agreed upon, to settle such a dispute, there is still no guarantee that parties only indirectly interested will carry out their mandate. Past experience does not encourage confidence in the reliability of such undertakings. And, most important of all, in addition to these negative failings that may be inherent in a plan devised to maintain peace, it may have the very serious positive defect of giving rise to a false sense of security. Here also past experience is not encouraging. Where there are brigands it is still the part of a wise traveler to carry arms.

Does this mean, then, that we are to regard as entirely hopeless all attempts to organize men's purposes to transcend personal interests, to work for greater and better harmony?

If, as statisticians, we have recognized the problem as being, in one respect at least, a problem of numbers, does this mean that the cards are stacked against us in our desire to establish and preserve a peace with honor and decency?

No, we shall not consider the case hopeless. Between an attitude of infantile optimism, at one extreme, and giving way to senile defeatism at the other, there surely is a middle ground of adult realism, which does not dwell in a world of illusions, but discerns difficulties for what they are.

When we labor to win the war, our aim is single and clear; the problem is only one of the means to achieve this clear purpose. But to speak of winning the peace is to telescope into one phrase what is really a discouragingly complex problem, even in merely focusing the target at which to aim.

The effort to win the war will be brought to a final climax in due time. But the peace can be secured and preserved only by lasting measures pursued unrelentingly through all time.

Among such measures may I in concluding suggest only one: A world cataclysm has welded together for the time being the United Nations. Let us work to preserve this unit. The parable of the mote and the beam applies to nations as it does to individuals. Let us bury minor criticisms of our Allies and friends in garlands of praise for the deathless glory of their magnificent stand against our common foe.

INTERNAL MIGRATION AND THE WAR*

By Henry S. Shryock, Jr.†
Bureau of the Census

During the past few years, workers and their families have migrated in such large numbers to centers of war production that they have created serious problems of housing, transportation, health, and education. These community stresses and strains have focused the public's attention on the population growth in many areas. There has been no systematic quantitative description, however, of the volume of internal migration or of the net gains made by the various areas of the country. Still less has the volume of recent migration been compared with the volumes of migrations in other periods of our history. The opinion has been expressed[1] that the extent of recent movements has been exaggerated in many of the popular stories in newspapers and magazines, since it takes a relatively small number of migrants to create numerous problems in a community, particularly an already congested one. There have been many attempts to attribute to population growth or in-migration, changes that may be due partly, or even solely, to fuller employment and greater purchasing power. Crowding on public vehicles; increased demand for gas, electricity, and telephones; larger payrolls or sales; and fewer housing vacancies have all been loosely interpreted as evidence of migration into an area. A better understanding of the problems before us should be afforded by isolation of internal migration, as one of the components.

The present consideration of internal migration is limited by the available data mainly to net migration for States and an approximation to net migration for counties. It would be extremely interesting to examine gross migration also, but pertinent statistics are exceedingly scanty. For other reasons, immigration and emigration to and from continental United States are not included. Movements of the latter type have by no means disappeared with the war, but they are relatively unimportant.

Estimates of net migration during the period April 1, 1940, to May 1, 1942, represent the difference between the estimated actual population on May 1, 1942, and the population estimated for that date on the assumption of no net movement. The estimates of actual population are those based mainly on the number of registrations for War Ration Book Number One (Sugar) and published in two recent releases of the Census Bureau.[2] The method of their preparation, together with brief comments on the limitations of the data, is described in these releases. A much more complete description of methodology and possible biases will be given in a forthcoming article by Norman Lawrence.[3] The population estimates omit persons in the armed forces, and the migration estimates are restricted to net interstate movements of civilians.

---

* This paper was prepared for presentation at the 104th Annual Meeting of the American Statistical Association which was to have been held at Cleveland, Ohio, December 29 to 31, 1942.

† For part of the discussion of metropolitan counties, the writer has drawn on an address read by Dr. Philip M. Hauser, Assistant Director of the Bureau of the Census, to the Wartime Marketing Conference of the American Management Association on January 15, 1943, in Chicago. He is also indebted to Mr. Robert J. Milliken of the Population Division of the Census Bureau, who supervised the calculations and gave valuable suggestions on methodology and interpretation.

[1] Myers, Howard B., "Defense Migration and Labor Supply," this Journal, Vol. 37, p. 72, March, 1942.

Obviously, these estimates of internal migration are affected by the possible biases elsewhere described in the rationing registration data as representative of the noninstitutional civilian population in an area. It is felt, however, that these shortcomings do not invalidate the general picture of internal migration here presented.

The other half of the comparison—the "expected" State civilian populations—was estimated by a variation of the standard method. From the population of a State at the census date, the births and deaths (both corrected for underregistration) through April, 1942, were added and subtracted, respectively. An estimate of the State's share of the net immigration from outside continental United States was also added. Finally, an allowance was made for the number of persons who had entered the armed forces from each State. The resultant estimate of expected civilian population subtracted from the corresponding actual population should represent the postcensal net civilian in-migration from or out-migration to other States. The algebraic sum of these figures for all States is zero.

For the United States as a whole, inductions and enlistments exceeded the natural increase combined with the small net civilian immigration from abroad, so that there was a slight loss in civilian population during these two years. The net migration figures presented in the table are roughly similar to the changes in civilian population. Whereever the estimated net migration is algebraically greater than the estimated population change, then the loss of men to the armed forces must have exceeded natural increase, and vice versa. Thus the high fertility States in the South show more increase in civilian population than in-migration or less decrease in population than out-migration.

Nineteen States and the District of Columbia had net in-migration between 1940 and 1942; 29 States had net out-migration. The geo-

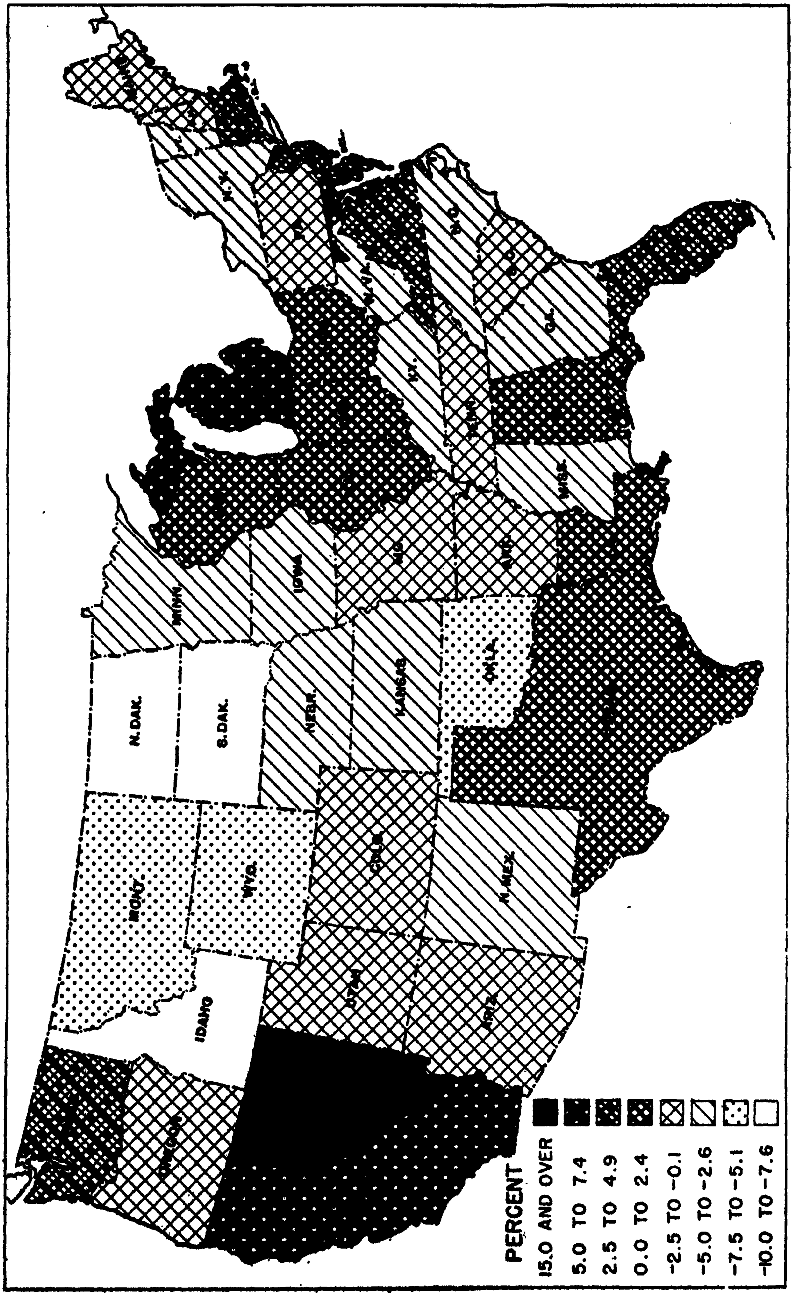
---

[2] U. S. Bureau of the Census. *Sixteenth Decennial Census of the United States.* "Preliminary Estimates of the Civilian Population by States: May 1, 1942," Series P-3, No. 30, October 27, 1942; and "Preliminary Estimates of Civilian Population in Selected Metropolitan Counties: May 1, 1942," Series P-3, No. 31, December 1, 1942.

[3] Entitled "War Registration as a Source of Population Estimates."

CHART I

NET CIVILIAN MIGRANTS 1940 TO 1942 AS PER CENT OF 1940 CIVILIAN POPULATION



graphic pattern may be seen from a glance at Chart I. In New England the rural, northern States lost civilian population by migration, and the urban, southern ones gained. Among the Middle Atlantic States, only New Jersey had a net in-migration. All of the East North Central States gained through migration, whereas all of the relatively more rural West North Central States lost. Losses were particularly severe in the tier of Great Plains States extending from North Dakota through Oklahoma. In the South, net in-migration was experienced by Delaware, Maryland, the District of Columbia, Virginia, Florida, Alabama, Louisiana, and Texas. These are all coastal States, and their gains were concentrated largely in their coastal metropolitan areas. All the Mountain States except Nevada suffered net out-migration. On the Pacific Coast, during the period under consideration, net gains occurred in California and Washington.

The highest percentage gain from net in-migration occurred in the Pacific Division, where there was a net gain amounting to 4.5 per cent of the 1940 civilian population. This was followed by the East North Central Division with 2.4 per cent. The New England and South Atlantic Divisions had net in-migrations of one per cent or less. The largest percentage of net out-migration (3.3 per cent) was suffered by the West North Central Division. Almost as high a net out-migration percentage was found in the Mountain Division, which was followed by the East South Central, Middle Atlantic, and West South Central Divisions, in that order. The largest number of net in-migrants—638,000 persons—was received by the East North Central Division; and the largest number of net out-migrants—448,000—left the West North Central Division.

California led the States in total number of net in-migrants with 395,000, and 4 States (Michigan, Illinois, Virginia, and New Jersey) and the District of Columbia had between 100,000 and 300,000. On the other hand, New York lost 421,000 through net out-migration, and 4 States (Oklahoma, North Carolina, Kentucky, and Minnesota) lost between 100,000 and 200,000. The District of Columbia's net migration was almost 25 per cent and far surpassed the percentage gain of any of the States. Only four of these (Nevada, California, Michigan, and Maryland) registered gains of five per cent or more. Material to be presented for counties will show how unevenly these migratory gains were distributed within the States having large net in-migration.

It should be emphasized before proceeding, however, that the figures considered above deal only with the balance of in-migration and out-migration for each State. The gross interstate migration was undoubtedly far larger. For example, according to provisional figures from

ESTIMATED NET CIVILIAN MIGRATION, 1940 TO 1942, 1930 TO 1940, AND 1920 TO 1930

Region, Division, and State	Net migrants, 1940 to 1942		Average number of migrants per year			Average yearly migration rate		
	Number	Per cent of 1940 population	1940 to 1942	1930 to 1940	1920 to 1930*	1940 to 1942	1930 to 1940	1920 to 1930*
Regions:								
The North	−122,000	−0.2	−58,700	−49,600	−4,300	−0.1	−0.1	—
The South	−182,000	−0.4	−87,300	−85,800	−133,400	−0.2	−0.2	−0.4
The West	304,000	2.2	146,000	135,400	137,700	1.1	1.1	1.3
The North:								
New England	80,000	1.0	38,400	−2,000	−31,400	0.5	—	−0.4
Middle Atlantic	−392,000	−1.4	−188,200	21,900	47,100	−0.7	0.1	0.2
East North Central	638,000	2.4	306,200	−1,800	67,000	1.1	—	0.3
West North Central	−448,000	−3.3	−215,100	−67,700	−87,000	−1.6	−0.5	−0.7
The South:								
South Atlantic	107,000	0.6	51,400	10,400	−65,400	0.3	0.1	−0.4
East South Central	−197,000	−1.8	−94,600	−42,600	−58,600	−0.9	−0.4	−0.6
West South Central	−92,000	−0.7	−44,100	−53,600	−9,400	−0.3	−0.4	−0.1
The West:								
Mountain	−132,000	−3.2	−63,300	5,100	−23,900	−1.6	0.1	−0.7
Pacific	436,000	4.5	209,300	130,300	161,600	2.1	1.5	2.3
New England:								
Maine	−16,000	−1.9	−7,700	−100	−5,700	−0.9	—	−0.7
New Hampshire	−7,000	−1.4	−3,400	1,300	−2,500	−0.7	0.3	−0.5
Vermont	−12,000	−3.4	−5,800	−1,900	−2,900	−1.7	−0.5	−0.8
Massachusetts	21,000	0.5	10,100	−6,700	−18,200	0.2	−0.2	−0.4
Rhode Island	22,000	3.1	10,600	100	−2,000	1.5	—	−0.3
Connecticut	72,000	4.2	34,600	5,300	−100	2.0	0.3	—
Middle Atlantic:								
New York	−421,000	−3.1	−202,100	49,200	52,200	−1.5	0.4	0.5
New Jersey	126,000	3.0	60,500	300	29,000	1.4	—	0.8
Pennsylvania	−97,000	−1.0	−46,600	−27,600	−34,100	−0.5	−0.3	−0.4
East North Central:								
Ohio	77,000	1.1	37,000	−4,500	13,900	0.5	−0.1	0.2
Indiana	76,000	2.2	36,500	2,400	300	1.1	0.1	—
Illinois	188,000	2.4	90,200	−4,600	21,400	1.1	−0.1	0.3
Michigan	282,000	5.4	135,300	4,600	40,700	2.5	0.1	0.9
Wisconsin	15,000	0.5	7,200	300	−9,300	0.2	—	−0.3
West North Central:								
Minnesota	−102,000	−3.7	−49,000	3,500	−19,900	−1.8	0.1	−0.8
Iowa	−89,000	−3.5	−42,700	−8,300	−20,500	−1.7	−0.3	−0.8
Missouri	−31,000	−0.8	−14,900	−6,100	−9,400	−0.4	−0.2	−0.3
North Dakota	−52,000	−8.1	−25,000	−11,400	−9,500	−4.1	−1.7	−1.5
South Dakota	−55,000	−8.6	−26,400	−11,300	−5,700	−4.3	−1.7	−0.8
Nebraska	−61,000	−4.6	−29,300	−15,800	−10,700	−2.3	−1.2	−0.8
Kansas	−58,000	−3.2	−27,800	−18,300	−11,300	−1.6	−1.0	−0.6
South Atlantic:								
Delaware	12,000	4.5	5,800	1,800	−2,200	2.1	0.7	−1.0
Maryland	93,000	5.1	44,600	9,500	−800	2.4	0.5	−0.1
District of Columbia	164,000	24.9	78,700	16,300	2,800	10.6	2.7	0.6
Virginia	130,000	4.9	62,400	−900	−23,600	2.3	—	−1.0
West Virginia	−60,000	−3.2	−28,800	−9,800	−6,400	−1.5	−0.5	−0.4
North Carolina	−168,000	−4.7	−80,600	−10,000	−1,000	−2.3	−0.3	—
South Carolina	−43,000	−2.3	−20,600	−12,700	−25,900	−1.1	−0.7	−1.5
Georgia	−85,000	−2.7	−40,800	−14,800	−41,900	−1.3	−0.5	−1.5
Florida	64,000	3.4	30,700	31,000	33,600	1.6	1.9	2.7
East South Central:								
Kentucky	−108,000	−3.8	−51,900	−11,600	−20,800	−1.9	−0.4	−0.8
Tennessee	−25,000	−0.9	−12,000	−3,100	−12,000	−0.4	−0.1	−0.5
Alabama	25,000	0.9	12,000	−19,800	−15,300	0.4	−0.7	−0.6
Mississippi	−89,000	−4.1	−42,700	−8,100	−10,500	−2.0	−0.4	−0.5
West South Central:								
Arkansas	−17,000	−0.9	−8,200	−17,300	−21,600	−0.4	−0.9	−1.2
Louisiana	46,000	2.0	22,100	1,700	−3,000	0.9	0.1	−0.2
Oklahoma	−168,000	−7.2	−80,600	−32,400	−5,600	−3.6	−1.4	−0.3
Texas	47,000	0.7	22,600	−5,600	20,800	0.4	−0.1	0.4
Mountain:								
Montana	−31,000	−5.6	−14,900	−1,300	−8,800	−2.8	−0.2	−1.6
Idaho	−50,000	−9.5	−24,000	3,400	−5,800	−4.8	0.7	−1.3
Wyoming	−13,000	−5.4	−6,200	400	−600	−2.6	0.2	−0.3
Colorado	−28,000	−2.5	−13,400	1,500	−3,700	−1.2	0.1	−0.4
New Mexico	−15,000	−2.8	−7,200	1,500	−2,500	−1.4	0.3	−0.6
Arizona	−10,000	−2.0	−4,800	500	1,300	−1.0	0.1	0.3
Utah	−6,000	−1.1	−2,900	−2,400	−4,000	−0.5	−0.5	−0.8
Nevada	21,000	19.1	10,100	1,500	200	8.5	1.5	0.2
Pacific:								
Washington	49,000	2.9	23,500	12,800	2,800	1.4	0.8	0.2
Oregon	−8,000	−0.7	−3,800	11,200	7,200	−0.4	1.1	0.8
California	395,000	5.7	189,600	106,300	151,600	2.7	1.7	3.2

from Thornthwaite, C. Warren, *Internal migration in the United States* (Study of Popu-
ution), Philadelphia, University of Pennsylvania Press, 1934, Map facing p. 22. Ex-
from abroad and migration of children born during the decade.

the 1940 Census, Georgia, which has been experiencing net out-migration for years, attracted over 150,000 in-migrants, 1935 to 1940.

The Bureau of Agricultural Economics estimates[4] that in 1941, when the net migration was 1,357,000 from farms, 2,171,000 persons moved from farms and 814,000 persons moved to farms.

Furthermore, more persons moved within States than moved from one State to another. According to the WPA surveys described by Myers,[5] the average move of workers to most defense centers during the pre-war phase of this period was less than 125 miles. Between 1935 and 1940, moreover, provisional census data indicate that in most of the Southern States the number of migrants from one county to another within the State exceeded the number from other States.

The effect of the defense and war programs on internal migration can best be evaluated by a comparison with internal migration during earlier periods. Unfortunately, no comparable figures are readily available for the last war or even for the entire decade 1910 to 1920. Estimates of net interstate migration from 1920 to 1930 were computed by Thornthwaite[6] on the basis of census age data. Similar estimates have been worked out for the decade from 1930 to 1940. (See table.) These net migration figures are for the total rather than the civilian population, but the armed forces in continental United States during these periods were negligible. All figures are for movements within continental United States, however. In order to allow for the different lengths of the periods, comparisons are made on the basis of the average number of net migrants per year. In the computation of annual net migration rates, the population at the mid-point of the period was used as the base.

It must be recognized that there was not a uniform annual flow of migrants during these periods; sometimes even the direction of net migration may have shifted from year to year. At the beginning of the 'twenties, adjustments were still being made to the population shifts effected by World War I; there was a severe but relatively brief depression followed by a long period of prosperity. The second of these decades contained years of severe depression and partial recovery.

Since there is more opportunity for the flow of migration to be reversed in a 10-year than in a 2-year period, average yearly net migration tends to be smaller in the longer periods. By and large, however, the

---

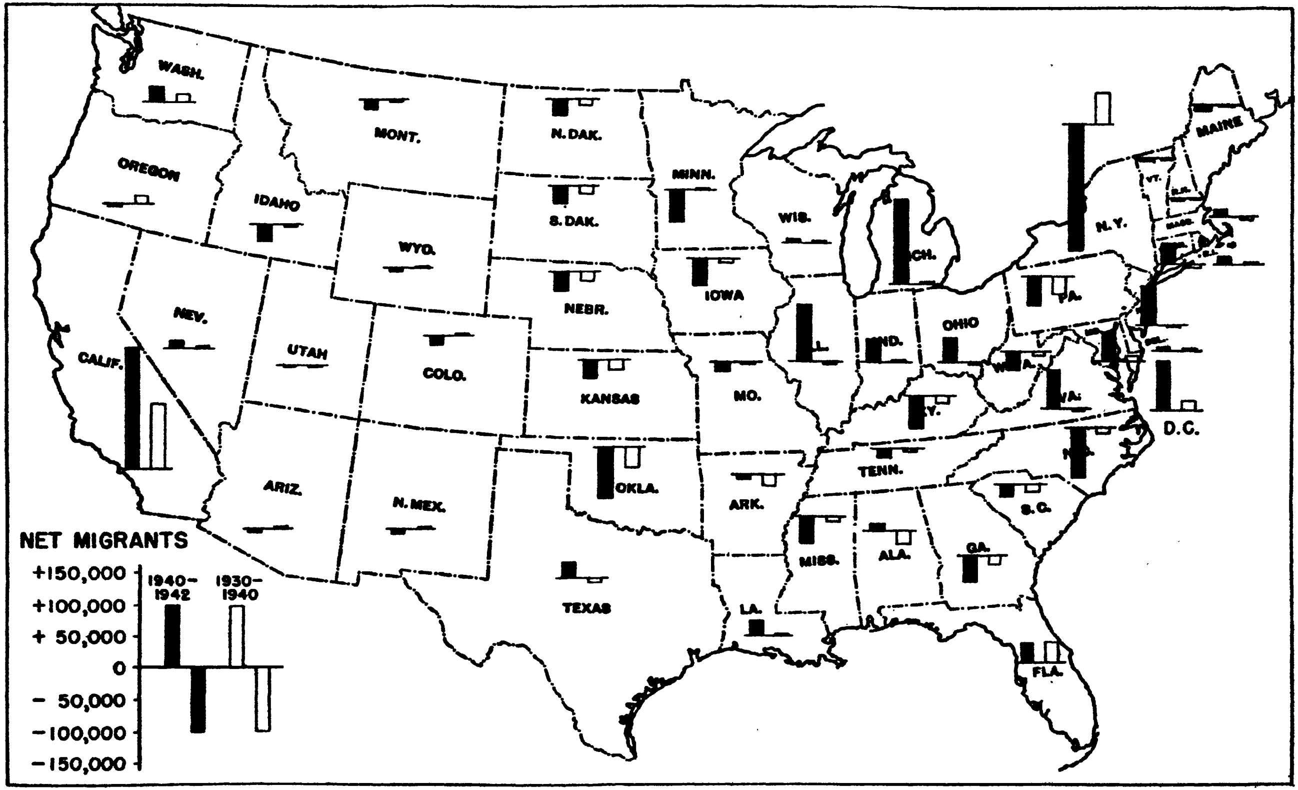
[4] Bureau of Agricultural Economics, "Farm Population Estimates: 1910–1942." Washington, November, 1942, p. 2.

[5] Myers, *op. cit.*, p. 74.

[6] Thornthwaite, C. Warren, *Internal Migration in the United States* (Study of Population Redistribution) Philadelphia, University of Pennsylvania Press, 1934, map facing page 22. Immigrants from abroad have been removed. These estimates, unlike those for the other periods, do not include movements of children born during the period.

CHART II

AVERAGE YEARLY NUMBER OF NET CIVILIAN MIGRANTS, 1940 TO 1942 AND 1930 TO 1940



direction of net migration for a State was probably fairly consistent within periods. It is pertinent to point out in this connection that, according to the Bureau of Agricultural Economics, there has been a net migration to farms only once (1932) since 1920 despite the sharp economic fluctuations in these years.

The Pacific Division was the only one that had a net in-migration during the postcensal years and in both of the preceding decades. The West North Central and East South Central Divisions had net out-migrations in all three periods, as did the West South Central except that in-migration almost balanced out-migration during the 'twenties. Net movements in the other divisions have not been so consistent.

Among the States, the balance of migration has consistently favored the District of Columbia, Florida, California, and, to a lesser degree, Nevada and Washington. The depression of the 'thirties interrupted the flow into industrial New Jersey, Ohio, Illinois, and Michigan. Many States had net out-migration in all three periods: Vermont, Pennsylvania, Iowa, Missouri, North and South Dakota, Nebraska, Kansas, West Virginia, South Carolina, Georgia, Kentucky, Mississippi, Arkansas, Oklahoma, Montana, and Utah. The war brought net in-migration for the first time in Alabama, whereas New York had not previously experienced a net out-migration.

A comparison of the yearly rates of net migration among the three periods (disregarding signs) shows that in most divisions and States net migration was larger in the war period than in the two preceding decades.[7] This was probably true also of gross internal migration. It is likely that migration from the Great Plains was even heavier than during the severe drought years of the 'thirties. The in-migration for Florida, however, has been declining.

An examination of internal migration in terms of States alone is not very satisfactory. Wide intrastate differentials in terms of subregions, urban and rural areas, and metropolitan districts and other territory, would certainly be expected. The Office of Price Administration in Washington has compiled the number of registrations for War Ration Book Number One by counties but not for cities or metropolitan districts. Estimates of civilian population on May 1, 1942, were accordingly prepared for counties. In order to obtain approximations to the census metropolitan districts, which are defined in terms of townships or other minor civil divisions, those counties with 50 per cent or more of their total 1940 population living in a metropolitan district were

[7] The rates for the period from 1920 to 1930 would have been slightly higher if the movement of children under 10 years old at the end of the decade had been included, so that comparisons with this decade must be made with caution. Usually, however, this age group is not highly mobile.

designated "metropolitan counties," and estimates were prepared for such counties, or groups of counties, corresponding roughly to each metropolitan district.

It would be too laborious a task to compute net civilian migration for each county by the method used for each State. Consequently, we shall have to infer migratory movements from population growth. It may be recalled that natural increase was about the same as inductions into and enlistments in the armed forces for the country as a whole. If this were true for each area, then any change in civilian population could be attributed to migration. Although there are geographic variations in age composition, sex ratio, fertility, and mortality (which would affect shifts of civilian population to the armed forces as well as natural increase), there is, for States, a high correlation between net migration and civilian-population change. Except for the influence of natural increase, the contrast between the growth of metropolitan counties and that of the rest of the country would be even greater, since natural increase is usually lowest in these counties.[8]

Of 137 metropolitan counties or groups of counties, 88 showed a gain in civilian population between April 1, 1940, and May 1, 1942, 43 showed a decline, and 6 showed no appreciable change.[9] For those counties in which the civilian population increased, the total gain amounted to 2,639,000, or 6.8 per cent of their 1940 civilian population. The net increase in the civilian population of all of the metropolitan counties was 1,853,000, or 2.8 per cent.

Very large gains in civilian population have taken place in the metropolitan counties centering in the following cities: Detroit, a gain of 336,000; Washington, D. C., 231,000; Chicago, 149,000; Los Angeles, 131,000; Norfolk, Portsmouth, and Newport News, 107,000; St. Louis, 97,000; San Diego, 97,000; and San Francisco and Oakland, 95,000. The largest percentage gains were recorded in the areas centering in San Diego; Norfolk, Portsmouth, and Newport News; Mobile; Corpus Christi; and Washington, D. C.—all of which had a civilian population gain of more than 25 per cent. The largest loss, 365,000, occurred in the New York-Northeastern New Jersey metropolitan counties.

Population changes in the metropolitan counties may be examined for four regions: the Northeastern, the North Central, the Southern, and the Western. The metropolitan counties in the South showed the greatest annual rate of increase in population between 1940 and 1942, 3.9 per cent, as compared with an annual rate of increase of 2.7 per cent in the West, 2.0 per cent in the North Central Region, and an annual rate of decrease of 0.6 per cent in the Northeast.

---

[8] The birth rate has risen markedly as a result of the war and its concomitant nearly full employment, but a comparison of natural increase with population increase in a sample of rapidly growing metropolitan districts shows that the excess of births over deaths was a minor factor.

[9] U. S. Bureau of the Census. *Sixteenth Decennial Census of the United States.* "Preliminary Estimates of Civilian Population in Selected Metropolitan Counties: May 1, 1942," Series P-3, No. 31, December 1, 1942.

Moreover, in the South the small metropolitan counties (those having central city populations of less than 100,000 or a total population of less than 150,000) had a slightly higher annual rate of increase than did the large metropolitan counties, 4.1 per cent as compared with 3.9 per cent. In each of the other three regions, the large metropolitan counties grew more rapidly than did the small.

As might be expected, metropolitan counties with high rates of growth were in practically every case those with more than average wartime activity, whereas those with heavy rates of decline were relatively inactive and contained relatively few important plants working on war contracts. There was no apparent relationship, however, between the industrial distribution of employed workers in a metropolitan area and its rate of growth, except that centers of shipbuilding and aircraft production had very large population increases.

Again we should compare the special type of population growth occurring in the defense and war periods with the growth of metropolitan counties during earlier years. The average yearly rate of increase from 1940 to 1942 was 1.3 per cent, which is significantly larger than the average rate of 0.8 per cent prevailing in the years from 1930 to 1940, which included a number of depression years. The recent rate falls short, however, of the yearly rate (2.7 per cent) prevailing between 1920 and 1930, a period when urbanization was extremely rapid. It was only in the South that the yearly rate of increase since the last census was greater than the yearly rate of increase both between 1930 and 1940 and between 1920 and 1930 (3.9 per cent as compared with 1.7 per cent and 2.9 per cent).

It is possible, however, that there was an even greater concentration of population during the defense period than during the 'twenties. Not only did metropolitan counties grow at the expense of other counties; but among metropolitan counties it has been shown that in most parts of the country, the larger ones grew more than the smaller.

Since the size of the civilian population of the country as a whole changed very little between April, 1940, and May, 1942, and since the 137 metropolitan counties, which contained about half the total population, had a gain of about 2.8 per cent, it is obvious that the remaining counties must have lost population at about this rate. In fact, only about one-fifth of these counties gained civilian population. Only 9.4 per cent gained as much as 5 per cent, only 2.9 per cent as much as 15 per cent, and only 1.3 per cent as much as 25 per cent. Many of the counties (37.9 per cent) gained or lost less than 5 per cent of their 1940 civilian population during these 25 months. More than half of the counties, 52.6 per cent, declined more than 5 per cent, 20.0 per cent declined more than 15 per cent, and 2.4 per cent declined more than 25 per cent.

Nonmetropolitan counties with extremely large increases in civilian population were mostly in the South and West, and many of them contain ordnance plants or army camps. (Civilians are attracted to the vicinity of army camps, first, by the need for construction workers and, later, by the opportunity to sell goods and services to soldiers.) Typical examples are Mineral and Clark Counties, Nevada, boom-centers of basic magnesium production, with a naval ammunition depot and an airfield, respectively; the independent city of Radford, Virginia, which contains a shell loading plant; Kitsap County, Washington, site of one of the country's largest navy yards; Orange County, Texas, a shipbuilding center; and Fairfax County, Virginia, which is partly in the Washington, D. C., metropolitan district but was not classified as a metropolitan county, according to the 50-per cent rule which was used.

In considering the net migration for States from 1940 to 1942, several States (Nevada, California, Michigan, Maryland, and Virginia) were singled out as having experienced very high rates of net in-migration. It is now possible to show how these gains were distributed within these States. In Nevada the population gain in Mineral County and in Clark County exceeded that in the State as a whole, indicating a loss in the balance of the State. Similarly, the combined gains in the Los Angeles, San Francisco, and San Diego metropolitan counties exceeded those in the State of California, although there were also sharp increases in a few of the small counties. In Michigan the Detroit metropolitan counties alone had a larger population increase than the State as a whole, but Kalamazoo, Muskegon, and a few other counties also shared in the gain. More than nine-tenths of Maryland's population increase occurred in the Baltimore and Washington metropolitan counties. Finally, Virginia's increase was concentrated in the Hampton Roads area, the Washington metropolitan counties, and in a few of the mountain counties where the increase generally was due to operation of a single large industrial plant. In short, even in these States a majority of the counties lost civilian population, largely through out-migration to other counties in the State.

The situation was more consistent in States with heavy out-migration. In the tier of States from North Dakota down through Oklahoma, all of which lost population, only 33 counties increased, and 4 of these were metropolitan counties.

The shifts of population since 1940 cannot be described simply as rural-to-urban migration. It has been shown that almost one-third of the metropolitan counties or groups of metropolitan counties lost civilian population. Many of these were in the Northeast. Counties in and around such large cities as New York, Pittsburgh, and Boston were included. It appears rather that movements were primarily to industrial areas active in war production or to scattered rural counties containing army camps or newly erected war plants. Although according to estimates by the Bureau of Agricultural Economics,[10] the net migration from farms in 1941 (1,357,000) was the largest on record, during this same year the WPA migration surveys showed that few workers migrating to defense centers came from agriculture. Possibly the migrants from farms filled the "vacuum" left by migrants to defense centers.

Since the characteristics of persons registering for War Ration Book Number One have not been compiled, the characteristics of net civilian migrants are not obtainable from estimates based on this source. The most extensive source of information on the characteristics of net migrants is represented by previously mentioned surveys conducted by the WPA. These were restricted to in-migrants into about 50 cities of 25,000 or more inhabitants and covered movements during a period of roughly one year, beginning in October, 1940. Unfortunately for our purposes, some of the more pertinent characteristics, such as age and sex, are shown only for workers and not for all in-migrants.

Statistics on color and size of family are among those shown for all in-migrants. The surveys also presented the origin of all in-migrants in terms of geographic area, distance, and size of place. Distributions presented only for migrant workers include those by age and sex, industry and occupation before and after migration, and employment status in present place of residence (by sex, age, month of arrival, and industry and occupation at last place of residence).

In summarizing these statistics, Myers states that in proportion to their number in the population, Negroes have made up only a small proportion of the migrants to defense centers. "In most areas fewer than five per cent of the migrants are Negroes, and even in the South, migration rates for Negroes are much lower than for whites. This is understandable, in view of the wide-spread discrimination against Negroes in war industries. It contrasts sharply with experience in the first World War, however, when a large scale migration of Negroes to Northern industrial centers took place. As the demand for labor increases and present employment restrictions are relaxed, it is probable that Negroes will begin to move in greater numbers."[11]

---

[10] *Loc. cit.*

[11] Myers, *op. cit.*, p. 74.

The migrant workers included relatively few females, again because of discrimination in hiring. Females were probably also underrepresented in the total number of migrants partly for this reason and partly because many male migrant workers have left their wives and other female dependents behind. In most defense centers, between 20 and 50 per cent of lone migrants had dependents living elsewhere. It is possible, of course, that the breadwinner, after a period of adjustment, was joined by other members of his family. Housing shortages would make the uniting of these families very difficult, however.

As usual, the migrant workers were characteristically young persons. "In nearly half of the cities the average age of all migrant workers is less than 30 years; in no city does the average rise as high as 35 years."[12] The average age of all migrants must have been even younger.

It is difficult to find comparable statistics for earlier migrations. The net interstate migration during the period 1930 to 1940 has been estimated by age, color, and sex, and a very rough comparison may be made on the basis of these data. It is not certain, of course, that the groups having the highest net migration rate also had the highest gross rate, but such a correspondence seems likely. During the last intercensal decade, the mobility of Negroes on the basis of net interstate migration was much higher than that of whites. Negro females were more mobile than Negro males, and white males were only slightly more mobile than white females. These differences seem to contrast sharply with those found in the early period of defense activity. Migrants in the two periods were probably more similar in terms of age composition. Between 1930 and 1940, persons 25 to 29 years old had larger net in- or out-migration rates than persons in any other 5-year age group. Since the migration occurred at an unspecified time during the decade, these young adults could have been up to 10 years younger when they made the indicated move from one State to another.

It is well to preface a summary by reminding the reader that accuracy of the net civilian interstate migrations from 1940 to 1942, which are the main feature of this article, depends on the accuracy of population estimates based on statistics of registration for sugar rationing, some of which figures are still provisional. It is clear, nonetheless, that the internal migration brought about by the war and defense programs is of a very different character from that of movements in the preceding years of peace. The special character and location of war jobs, the withdrawal of men from the civilian labor force into the armed forces, and the rapid tempo of economic change have produced a very unusual situation. This situation is reflected in the relatively limited number of areas that have attracted large numbers of in-migrants and in the personal characteristics of the in-migrants.

---

[12] Myers, *op. cit.*, p. 75.

Other data also confirm the conclusion that internal migration in the defense and early war periods can not be characterized simply as a rural-to-urban movement. Those metropolitan counties and a few scattered rural counties that have been active in war production because of the contracts secured have been practically the only areas to gain civilian population by net migration. In some cases where a large army camp or ordnance plant had to be built in jig time in the open country, the population during the period of construction may have exceeded the present population. Population growth, and consequently net in-migration, have been phenomenal in some areas. In other areas where extraordinary growth has been claimed in overenthusiastic chamber of commerce and newspaper estimates, net in-migration has been small or nonexistent.

During the period covered by the WPA migration surveys, from October, 1940, to the summer or fall of 1941, it was observed that in-migration to defense centers was proceeding at an accelerating pace. Estimates of population made for Washington, D. C.,[13] in the three six-month periods from July 1, 1940, to January 1, 1942, show successively larger rates of population increase. There is also evidence of an acceleration of migration since May 1, 1942, into certain areas, such as the Baltimore and Portland, Oregon, metropolitan counties. For the country as a whole, it can not be ascertained whether the peak of war-induced internal migration has yet been reached or, if so, when.

In addition to the usual selection of young adults in the wartime migration, there seem to have been, at least in the first year or two of the postcensal period, selections on the basis of color and sex. White males seemed to have formed a larger proportion of the migrant workers to defense centers than would be expected on the basis of their proportion in the labor force. It is possible, however, that their places may have been taken by a secondary migration of other persons. As the war proceeds and the manpower situation becomes even more serious, there may be waves of migration of Negroes, women, and even elderly persons. If migrant husbands are eventually able to find living accommodations for their families, the sex ratio among migrants will be reduced by this cause alone.

---

[13] U. S. Bureau of the Census. *Sixteenth Decennial Census of the United States.* "Estimated Population of Washington, D. C., by Six-Month Periods: July 1, 1940, to January 1, 1942," Series P-3, No. 27, May 25, 1942.

According to Myers, the defense migration has been remarkably successful. Unemployment among migrant workers in defense centers is very low, and it is probable that most of them have improved their economic position. The wartime internal migration, however, has created many problems, both for the migrants and for the communities that they have entered. There are obviously many new problems connected with local government, health, housing, traffic, and education. The process of migration and conditions in boom towns have had important repercussions upon the family, juvenile delinquency, and recreation. People who are interested in the decentralization of industry as a means of securing more wholesome living conditions and a more nearly uniform distribution of economic well-being are now confronted by greater concentration of industry and population than existed in 1940. Even before the war, students like Creamer[14] have shown that decentralization of industry was not going on.

Hauser, in an address on "Population Shifts and Income Changes" (as yet unpublished), read before the American Management Association on January 15, used the quartile in which the rate of growth of a metropolitan county or group of counties fell during each period to evaluate the permanency of the wartime population shifts. (The periods were the same as those used in this article; namely, 1940 to 1942, 1930 to 1940, and 1920 to 1930.) On the basis of the simple criterion of previously observed growth during two periods characterized by radically different economic conditions, he concluded that on "a regional basis it is clear that the metropolitan centers which have grown most rapidly and which . . . have the best prospect of retaining their growth are in the South and West."

Whether the enormous wartime gains in population made by particular areas are retained after the war would seem to depend in large measure upon the resiliency of the area's economy. If the industries in these areas can not shift readily to peacetime production, there may well be wholesale repatriation of migrants, leaving ghost towns in their wake, or else depressed areas supporting heavy relief loads. This last would be further complicated by residence requirements for relief and unemployment compensation. The major determinant in the post-war period, however, may be drastically new government policies and programs that will far outweigh the effects of the factors just mentioned.

---

[14] Creamer, Daniel B., *Is Industry Decentralizing?* (Study of Population Redistribution), Philadelphia, University of Pennsylvania Press, 1935.

POPULATION AND WAR LABOR SUPPLY*

By William H. Mautz and John D. Durand
Bureau of the Census

The effectiveness with which a modern nation can wage total war is primarily a function of manpower, materials, and management. During our first year in the war, we have made great progress in solving problems of materials and management; manpower is now the Number 1 problem. The length of the war will depend to a large extent upon our success in attaining the most effective utilization of this nation's human resources.

Labor supply problems are involved in nearly every phase of the war effort. They include allocation of manpower between the armed forces and civilian production, transfer of workers from nonessential industries, reduction of unemployment and absenteeism, lengthening working hours, more efficient utilization of labor resources by means of training and upgrading, and increasing the total labor supply mainly by recruiting additional workers from the population normally not in the labor force. This paper deals with the latter part of the problem; the mobilization of the nation's potential labor supply.

The War Manpower Commission's estimates of labor requirements call for a total labor force of 62,500,000 by December, 1943, including armed forces of 10.7 millions.[1] This means an increase during the current year of 3,500,000 over the estimated 59,000,000 in the labor force in December, 1942, and a total increase of 9,500,000 since March, 1940. From what sources in the population have the 6,000,000 workers come who were added to the labor force between March, 1940, and December, 1942? Can these sources supply another 3,500,000 workers in 1943? Upon what groups of the population and in what areas should the efforts to recruit additional workers be focussed?

These are primarily demographic questions, since the size, geographical distribution, and characteristics of the population are basic determinants of the nation's potential manpower and of the measures that are appropriate to realize it. Fortunately, the most comprehensive inventory of this nation's human resources that has ever been available is assembled in the tabulations of the 1940 population census, including a wealth of information about the characteristics and geographical distribution of the labor force and of persons not in the labor force. In addition, the monthly sample survey of the labor market (transferred to the Census Bureau from the Work Projects Administration in August, 1942) provides current data on the size and composition of the labor force.

---

* This paper was prepared for presentation at the 104th Annual Meeting of the American Statistical Association which was to have been held at Cleveland, Ohio, December 29 to 31, 1942.

[1] War Manpower Commission, estimates prepared in January, 1943.

SOURCES OF ADDITIONAL WORKERS

Since the first determinant of national labor supply is the size of the population, population growth is an important source of additional manpower. Growth of the population of working age provided almost 1,500,000 of the 6,000,000 increase in the labor force between March, 1940, and December, 1942, and will provide a further increase of 500,000 by December, 1943. The remaining 3,000,000 additional workers needed by the end of 1943 will have to be obtained by still more extensive employment of persons normally not in the labor force.

Employment of men.—There is very little room for expansion of the male labor force except among boys of school age. The 1940 census showed only 10,600,000 males 14 years old and over not in the labor force, of whom 3,700,000 were reported as unable to work or confined to institutions. The age classification of men not in the labor force in March, 1940, is shown in Table I.

TABLE I

MALES NOT IN LABOR FORCE: MARCH, 1940[*]

(millions)

Age	Total	Unable to work or in institutions	In school	Others not in labor force
14 years and over	10.6	3.7	4.6	2.3
14 to 17 years	4.0	0.1	3.5	0.4
18 to 44 years	2.5	0.8	1.1	0.7
45 to 64 years	1.5	1.0	—	0.5
65 years and over	2.6	1.8	—	0.7

[*] U. S. Bureau of the Census, "Labor Reserves in the United States by Age, Marital Status, and Sex" (Release Series P-14, No. 4), October 9, 1942. More detailed data will appear in the forthcoming Census Bulletin, *Characteristics of Persons not in the Labor Force.*

The 1,800,000 able-bodied men 18 to 44 years old who were not in the labor force at the time of the census were mostly older students, many of whom have been inducted into the armed forces since that time, or have left school early to take war jobs. Probably not very many more workers can be obtained from this group.

Most of the 1,200,000 men over 45 who were not in the labor force but able to work in 1940 have doubtless become workers since that time, and even the group of 2,800,000 shown in the census as unable to work or in institutions has apparently made a sizeable contribution to the labor supply. About 1,300,000 men over 45 were added to the labor force between March, 1940, and October, 1942. About 500,000 of these[2] were men over 65. These increases have resulted partly from population growth, partly from the postponement of retirements, and partly from the return to work of men who had previously withdrawn from the labor force.

There is still opportunity for some further expansion of the labor force in this direction. The monthly survey of the labor market for November, 1942, which included the question, "Could this person take a full-time job for wages if one were available within 30 days?" indicated that about 300,000 men over 45 years old would be available for work. About 200,000 of these were over 65 years of age.[3] Nevertheless, the increase in the labor force in the male population over 45 is nearing the limit imposed by disability, chronic illness, and old age. This part of the population cannot supply a large share of the additional labor force needed during the war.

Employment of children.—Boys and girls of 'teen age have provided a considerable part of the additional labor force up to the present. The usual influx of students into the labor market during summer vacations has been abnormally large in the last two years, and has constituted an important source of labor supply, especially in farming areas. In addition, year-round increases have been obtained from students dropping out of school to take war jobs, or doing part-time work after school hours. In March, 1940, the number of persons 14 to 19 years old in the labor force was 3,900,000;[4] the corresponding figure for March, 1942, was 4,800,000.[5] In the summer of 1942, seasonal employment of young people on a scale unprecedented in recent years raised the labor force in this age group to a peak of 8,400,000 in July. By October, 1942, the figure had declined seasonally to 5,600,000, but was still 1,700,000 higher than that for March, 1940. Much of this increase undoubtedly came from persons past high-school age, but a considerable part of it must have resulted from increased employment[6] of children under 18.

---

[2] U. S. Bureau of the Census, "Age of the Labor Force, by Sex, for Regions, Divisions, and States, and for Cities of 100,000 or More: 1940" (Release Series P-16, No. 1), August 20, 1942; and *Monthly Report on the Labor Force*, "Age and Sex of Persons in the Labor Force: June 1941–October 1942" (MRLF-No. 4), December 12, 1942. Although the results of the Work Projects Administration's "Sample Monthly Report of Unemployment" for the week of March 24 to 30, 1940, were not in complete agreement with the employment status data obtained in the decennial census enumeration for the same week, it is believed that the differences do not invalidate any of the conclusions of this paper. Studies are in progress which will provide the basis for more precise comparisons between 1940 census data and the results of the monthly sample surveys.

[3] U. S. Bureau of the Census, *Monthly Report on the Labor Force*, "The Nation's Available Labor Reserve; November, 1942" (MRLF—No. 7). This report gives the distribution by sex, age, farm residence, previous work experience, and other characteristics of persons not in the labor force who indicated that they were available for full-time employment.

[4] U. S. Bureau of the Census, Release Series P-16, No. 1.

[5] U. S. Bureau of the Census, MRLF—No. 4.

The relatively high wages now being offered to inexperienced young workers will probably continue to attract young people from the schools to the war factories.[7] Full-time employment of children under 18 is restricted by child-labor and compulsory school-attendance laws, and by established standards of education, but these restrictions have been relaxed somewhat since this country entered the war. Some States have modified their child-labor laws, and there has been a tendency to enforce these laws less strictly in view of the manpower shortage. Not very many more children under 18, however, can be brought into the labor market as full-time workers unless a wholesale sacrifice of educational standards is decided upon for the sake of additional labor supply during the war emergency. Otherwise, the principal contribution of children under 18 to the war labor supply is likely to be in the form of part-time and seasonal work.

Employment of women.—Further large additions to the labor force will evidently have to come mainly from the female population. Women who enter the labor force normally remain for only a relatively short period after leaving school, and before getting married and starting to raise a family.[8]

During the last two years, women have been entering the labor market earlier and in larger numbers, and remaining longer than in normal times. Some women who had retired from the labor force to take up homemaking have returned to work, and a considerable number of housewives who were not previously workers have been attracted to the labor market by higher wages, better jobs, and patriotic appeals. Inductions of married men in the armed forces have forced some wives to take jobs to support themselves and their families. The absorption of women into the war labor force has been facilitated by establishing training schools for war production jobs, by simplifying machinery and breaking down complex industrial operations, and more recently by encouraging part-time work. Nurseries for the children of working mothers have been organized in many localities to free women for war work. The result of these developments has been an increase of 2,700,000 in the female[9] labor force between March, 1940, and December, 1942. Most of the increase has come from girls in their 'teens and early 20's, and women between the ages of 35 and 55.

---

[6] In the case of males, the majority of the increase between March, 1940, and October, 1942, for the age group 14 to 19 must have come from boys under 18, since the increase for the ages 14 to 19 (1,000,000) exceeded the number of males 18 and 19 years old not in the labor force in March, 1940 (850,000). The authors estimate that perhaps 800,000 of the total increase of 1,700,000 for males and females 14 to 19 years old occurred in the ages 14 to 17.

[7] In November, 1942, there were 500,000 girls and 200,000 boys 14 to 19 years old not in labor force who were available for full-time jobs. U. S. Bureau of the Census, MRLF—No. 7.

[8] For data on labor-force status of women by age, marital status, and household relationships, see: U. S. Bureau of the Census, "Household Relationships of the Labor Force" (Release Series P-14). More detailed statistics will appear in the forthcoming Census bulletin, *The Labor Force—Employment and Personal Characteristics.*

The authors estimate that the reservoir of women 18 years old and over not in the labor force stood at 33 million at the end of 1942. From this reserve must be drawn most of the 3,000,000 additional workers needed by the end of 1943. In November, 1942, 4,500,000 women not in the labor force were available for full-time employment.[10]

TABLE II

FEMALE POPULATION AND PER CENT IN THE LABOR FORCE, BY FAMILY STATUS; MARCH, 1940*

Family status	18 to 44 years old	18 to 24 years old	25 to 34 years old	35 to 44 years old
Female population (millions)				
Total	28.2	8.4	10.8	9.1
Married, husband present	18.2	3.3	7.9	7.0
Without children under 10	8.3	1.5	2.8	4.0
With 1 or more children under 10	9.9	1.9	5.1	2.9
Other marital status	10.0	5.0	2.9	2.1
Single	7.7	4.7	2.0	0.9
Widowed, divorced, or married, husband absent	2.4	0.3	0.9	1.2
Per cent in labor force				
All females	33.9	43.2	32.7	26.7
Married, husband present	16.0	15.6	17.4	14.6
Without children under 10	25.7	27.3	33.4	19.7
With 1 or more children under 10	8.0	6.4	8.8	7.6
Other marital status	66.3	61.5	74.5	66.5
Single	67.7	62.0	78.5	73.3
Widowed, divorced, or married, husband absent	61.4	53.9	64.9	61.1

* U. S. Bureau of the Census, Release Series P-18, No. 13.

Employment of women under 45.—The chief obstacle to the expansion of the labor force among women between the ages of 18 and 44 is the responsibility of caring for their homes and families, and especially for young children. The importance of these factors is shown by Table II on the labor-force status of women 18 to 44 years old according to family relationships, based on 1940 census data.

---

[9] U. S. Bureau of the Census, *Monthly Report on the Labor Force*, "Estimates of Labor Force, Employment, and Unemployment in the United States During the Week of December 6–12, 1942" (MRLF—No. 6), January 9, 1943.

[10] U. S. Bureau of the Census, MRLF—No. 7, special report on available labor reserves, November 1942.

Even in 1940, two-thirds of the unmarried women in this age group, and one-fourth of the married women without children were in the labor force. Only 8 per cent of the mothers of young children were workers, and these few were probably mostly women who could hire servants or arrange for someone to take care of the children during working hours, or mothers who were forced to work by pressure of poverty.

The distribution by family status of women 18 to 44 years old who were not in the labor force in 1940 shows that there was not very much room for expansion of the labor force in this age class without the employment of women with considerable home responsibilities. Of these women, 82 per cent were married with husband present in the household, and 60 per cent of those who were married had one or more children under 10 years of age. Only 18 per cent were single, widowed, divorced, or married, husband absent.[11]

Up to the fall of 1942, nearly 1,400,000 women 18 to 44 years old had been added to the 1940 labor force; about half of these were in the age group 18 to 24, and half in the group 35 to 44 years old.[12] Since most of these additions were probably drawn from the unmarried women and from the married women without children, the margin for further increase in the labor force through employment of women with relatively light home-making responsibilities is probably considerably narrower now than in 1940.[13] Moreover, substantial increases in birth and marriage rates since 1940 have still further reduced the reserve of unmarried women without children.[14]

In order to provide additional information regarding the relation between employment of women and their family responsibilities, the Bureau of the Census has made extensive tabulations of 1940 census data on labor-force status of women according to number of children and various social and economic characteristics. In these tabulations, women in and outside of the labor force are classified by age, marital status, number of children under 5 and 5 to 9 years old, color, value of home, income of husband, education, and other characteristics.[15]

---

[11] U. S. Bureau of the Census, Release Series P-18, No. 13.

[12] The October, 1942 *Monthly Report on the Labor Force* (MRLF—No. 4) showed an increase of 1,600,000 since March, 1940 for the female labor force in the age group 14 to 44 years. The authors estimate that about 200,000 of this increase occurred in the ages 14 to 17.

[13] Forty per cent of the women who reported that they were available for employment in November, 1942, stated that they were responsible for the care of children under 14 years of age. U. S. Bureau of the Census, MRLF—No. 7.

[14] Hauser, Philip M., "Population and Vital Phenomena," *American Journal of Sociology*, November, 1942.

[15] For results of these tabulations, see: U. S. Bureau of the Census, Release Series P-18, Nos. 1–13. Additional data will appear in the forthcoming Census Bulletins, *The Labor Force—Sample Statistics, Employment and Family Characteristics of Women*, and *Families—Employment Status*.

Employment of older women.—Among women over 45, family responsibilities are only a secondary factor limiting the expansion of the labor force, since comparatively few women over that age have any young children. Many of these women, of course, are so heavily burdened with keeping house for older children or other dependents, or for a large family of adults, that they have no time for other employment. These limitations, however, are probably less important than the handicaps of age and lack of recent work experience, which make it more difficult for women over 45 to get jobs than for younger women.

In spite of these handicaps, the number of women over 45 in the labor force has increased considerably during the last two years. Between March, 1940, and October, 1942, the female labor force increased by one-sixth in the age group 45 to 54 years; and by one-fifth in the group 55 to 64 years.[16] There has undoubtedly been a tendency for women reaching the ages at which they would normally retire, to postpone their withdrawal from the labor market; but this tendency alone is not sufficient to account for the large increases which have occurred. Employers have apparently had to hire more older women because younger workers were not available.

Women over 45 will probably continue throughout the war period to be a source of considerable numbers of workers. The 4,500,000 women not in the labor force but immediately available for work in November, 1942, included 1,300,000 women 45 to 64 years old and 100,000 over 65.[17] Unless the problem of relieving younger women of their family responsibilities is successfully solved, these older women will have to provide an increasing share of the needed labor supply. Training for older women who have not been employed for many years is especially important for the efficient utilization of this part of the potential labor supply.

The principal sources of additional labor supply in the population, and the rates at which these reserves are being absorbed into the labor force are summarized in Table III.

The increase in the employment of men over 45 has probably about exhausted the possibilities of labor-force expansion in this part of the population. In October, 1942, the proportion of workers had reached 92 per cent among men 45 to 64 years old, and nearly 50 per cent among those 65 years old and over. These proportions are probably approaching the limits imposed by old age and disability. The proportion of children 14 to 17 years old in the labor force increased from about one-eighth in March, 1940, to about one-fifth in October, 1942. Further increases in the employment of children will probably have to be confined largely to seasonal and part-time work. Increases in the employment of women have been much smaller in proportion to the size of this part of the potential labor supply. Women 18 to 64 years old, of whom there were over 28 million not in the labor force in October, 1942, are the population group upon which attention must be concentrated in the effort to add 3,000,000 more workers to the labor force by the end of 1943.

---

[16] U. S. Bureau of the Census, MRLF—No. 4.

[17] U. S. Bureau of the Census, MRLF—No. 7.

TABLE III

Population group	Labor-force increase March, 1940, to October, 1942 (millions)	Persons not in labor force, October, 1942* (millions)
Children 14 to 17 years old	0.8	7.4
Men 45 years old and over	1.3	3.6
Women 18 to 44 years old	1.4	18.1
Women 45 to 64 years old	0.5	10.4
Women 65 years old and over	—	4.6

* Estimated on the basis of data from the *Monthly Report on the Labor Force* for October, 1942 (MRLF-No. 4) and estimates of population by age and sex as of that date.

GEOGRAPHICAL DISTRIBUTION OF THE POTENTIAL LABOR SUPPLY

When expressed in national totals, the manpower problem is oversimplified. The concrete problem is to provide the labor supply needed in the aircraft factories and shipyards of Baltimore, on the farms of Aroostook County, Maine, and in each of hundreds of other local labor markets. Comparisons of national labor requirements with sources of additional workers in the national population are misleading unless allowance is made for the heavy concentration of manpower needs in certain areas.

Lack of balance between the geographical distribution of the population and the distribution of war labor requirements enormously complicates the problem of efficiently utilizing the nation's human resources. In areas where comparatively little essential war production is carried on, labor surpluses tend to develop when nonessential production is curtailed, and it is not easy to transfer these surpluses to areas of labor shortage. In the New York City area, for example, where there has been relatively little expansion of war production, labor reserves are not easily mobilized for employment in critical labor-shortage areas in upstate New York or even in nearby New Jersey.

Internal migration during the last two years has played a very important part in bringing the supplies of labor more nearly into balance with the demands.[18] The possibilities of migration are limited, however, by lack of housing, transportation, and other facilities, in the crowded cities where war industry is centered. In these areas, future increases in the labor force will have to be obtained to a greater extent from the population already present. Reserve labor supplies among women and young persons will have to be employed much more extensively in such areas than in the nation as a whole.

The provision of an adequate farm labor supply is one of the most difficult problems in the geographical distribution of manpower. Occupational deferments for agricultural workers will do much to halt the depletion of farm labor supplies through inductions into the armed forces, but the drain resulting from migration of workers from farms to industrial centers is likely to continue. Farm youths are attracted to the cities not only by prospects of higher wages but also by the more varied amusements and less isolated existence which urban life offers. Meanwhile, heavy demands are being made upon agriculture to provide food for the armed forces and the lend-lease program. The problem of agricultural labor supply is complicated by seasonal variations in farm labor requirements, by difficulties resulting from the remote situation of farms, and by special obstacles to the employment of women in the farm population.

Seasonal peaks in farm labor needs are normally met partly by extensive employment of the women and children in the farm population, and partly by means of hired labor, which is supplied in some areas largely by migratory workers. During the last two years the normal sources of both hired labor and family workers have been considerably reduced. The deficiencies have been met partly by still more extensive employment of women and children at seasonal peaks, and in some cases partly by organized recruiting of school children and women in nearby cities and towns to act as temporary workers for such operations as fruit harvesting. Even with the use of such devices it will probably be difficult to obtain the labor supply needed for the expanded 1943 agricultural production program.

The problem of replenishing the full-time farm labor supply is still more difficult. Urban workers are likely to be reluctant to take jobs on farms, because of the disadvantages of isolation, the arduous character of the work, and the comparatively low wages. Less than one-sixth of the persons not in the labor force who were available for jobs in November, 1942, were willing to take a farm job, although nearly one-fifth of them lived on farms.[19]

---

[18] Shryock, Henry S., Jr., "Internal Migration and the War," in this issue of the JOURNAL.

[19] U. S. Bureau of the Census, MRLF—No. 7.

Some of the year-round labor requirements can be supplied by farm housewives and by farm children working part-time after school, but there are many kinds of farm work that women and children cannot regularly do. Moreover, farm housewives tend to be heavily burdened with housekeeping duties which greatly limit their time for other work. The average wife in a rural-farm family has more children to care for than the city wife, and enjoys fewer of the conveniences which lighten the housekeeping load, such as washing machines and vacuum cleaners.

These obstacles are reflected in extremely low proportions of farm women in the labor force. Only 15 per cent of all women 18 to 44 years old in rural-farm areas (outside of metropolitan districts of 100,000 or more) were in the labor force in March, 1940; the corresponding proportion for women in other areas was 38 per cent. For married women, husband present, with one or more children under 10 years old, the proportions in the labor force were 4 per cent in farm areas and 9 per cent in other areas.[20]

Even during the war emergency, it will be very difficult to surmount these obstacles to the employment of women on farms. It seems probable that the contribution of farm housewives to the war labor supply will be mostly in the form of seasonal work and relatively light part-time chores at other times of the year.

The 1940 Census tabulations provide a large amount of information relating to potential labor supplies in particular areas. The tabulations mentioned above for women in and outside of the labor force by number of children, age, and other characteristics have been made for each metropolitan district of 100,000 inhabitants or more, and for urban, rural-nonfarm, and rural-farm areas of State groups.[21] In addition, statistics for persons not in the labor force, by sex, subdivided into those engaged in own home housework, in school, unable to work, in institutions, and others not in the labor force, have been tabulated for rural-farm and rural-nonfarm parts of counties and for each urban place.[22]

The 1940 data for particular labor-market areas can be applied to the analysis of present problems of labor supply if allowances are made for changes since the census date due to increases in employment, population growth, migration, and changes in marriage and birth rates. In most cases migration has probably had relatively little effect upon the reserves of potential workers not in the labor force, since the available evidence indicates that most of the migrants into centers of war production were in the labor force.[23]

---

[20] U. S. Bureau of the Census, Release Series P-18, No. 13.

[21] References to publications of these data are given above.

[22] U. S. Bureau of the Census, *Sixteenth Decennial Census of the United States: 1940, Population*, Vol. II, "Characteristics of the Population."

The problems of recruiting 3,000,000 more workers during 1943 from the population normally not in the labor force, and of training and placing them in war jobs, will tax the facilities and ingenuity of our manpower agencies. The nonworker groups least affected by family responsibilities, old age, physical disabilities, and other factors restricting employment, have for the most part already been added to the labor force. We are now faced with the task of taking an additional quarter of a million workers each month out of a labor reserve from which the most readily employable groups have already been taken. Only a small margin remains for further increase in the adult male labor force. A considerable reserve of seasonal workers exists among children under 18 years old, but only a small proportion of them will be available for full-time work the year-round, unless we are willing to make serious sacrifices in educational and child-labor standards. Home responsibilities, and particularly the necessity of caring for their children, restrict the employment of young women. Up to the present, additions to the female labor supply in the age range 18 to 44 years have probably come largely from unmarried women and married women with few homemaking duties—nonworker groups that were already comparatively small in 1940 and have dwindled steadily since that time. An important share of the additional war jobs has been filled by women between 45 and 65, comparatively few of whom have young children to care for, but such women are handicapped by age and by lack of recent experience, and many of them have heavy responsibilities for the care of older children or of large families of adults.

Married women 18 to 64 years old, without children under 10, are the most important population group still not in the labor force. This pool of reserve labor supply will probably have to be drawn upon more heavily in 1943, in spite of their family responsibilities. Inductions of married men without children are likely to bring a considerable number of wives into the labor force, especially from the group who have hitherto refrained from taking jobs because of the possible effect upon their husbands' draft classifications. Other factors which may lead to an increase in employment of married women are heavier tax burdens, rising cost of living, higher wages, and more attractive employment opportunities. It is not likely, however, that such factors alone will bring about the needed increase in the labor supply.

---

[23] Myers, Howard B., "Defense Migration," in: *The Manpower of the Nation in War Production*, Book One (Hearings before the Select Committee Investigating National Defense Migration, House of Representatives, 77th Congress, 2nd Session, H. Res. 113, Part 27), pp. 10318–10330, esp. p. 10325.

The employment of women with minor children, particularly children under 10 years old, should certainly not be encouraged if an adequate labor supply can be obtained from other sources. In the crowded centers of war industry, however, it is becoming more and more important to mobilize all possible labor reserves in the resident population, including some mothers of young children. If these women are to be employed, fairly extensive provisions for child care are likely to be needed in critical labor supply areas. To the extent that such measures succeed in cutting down further migration into these areas, they will greatly simplify problems of war housing, transportation, and post-war adjustment.

A national view of labor reserves in relation to manpower requirements gives only the general dimensions of the problem. National figures showing from what population groups additional workers can most easily be drawn, and what factors restrict their employment, are useful only as guides for establishing broad national policies. The crux of the manpower problem is in the local community. It is there that the problems of labor requirements, recruitment, training, and placement originate and must ultimately be solved. The success of national manpower policy will depend chiefly upon effective direction and coordination of the activities of national manpower agencies, regional and local War Manpower Committees, and joint labor-management committees in the hundreds of local labor markets throughout the country. Several months ago the President aptly summarized the problem with the statement, "It is not that we do not have enough people in this country to do the job. The problem is to have the right numbers of the right people in the right places at the right time."

STATISTICAL LAWS OF FAMILY EXPENDITURE*

By Holbrook Working
Food Research Institute, Stanford University

The first result of a statistical study is a body of data exhibiting certain facts regarding either a particular population or a particular sample from some population. Analysis of the statistics in conjunction with other pertinent information usually permits developing the original conclusions into broader statements of tendencies. If the indicated tendencies are sufficiently general, they may deserve characterization as scientific laws.

Statistical data on family expenditures have been collected in a number of countries at intervals over a long period of years. The most recent large surveys of family expenditures, and in many respects the most significant among all such surveys, are two which were made in the United States in the middle 1930's. The smaller of the two was conducted by the Bureau of Labor Statistics and was confined to families of wage earners and clerical workers. In that survey, data were obtained from some 16,000 families in 42 medium-sized and large cities and 13 small cities. The results have been published in a series of 8 volumes of the Bureau of Labor Statistics. The larger of the two surveys was a joint enterprise of the Works Progress Administration, the Bureau of Labor Statistics, and the Bureau of Home Economics, with the cooperation of the National Resources Planning Board and the Central Statistical Board, and covered more than 300,000 families of nearly all occupations and degrees of social status in all the major types of community in the United States. The results have been published in some 47 volumes issued by the Bureau of Labor Statistics (22 volumes), the Bureau of Home Economics (22 volumes), and the National Resources Planning Board (3 volumes).

The present paper gives some of the results of an attempt to draw from this great body of data some significant generalizations which may deserve to be regarded as laws of family expenditure.

The distribution of family expenditure is conditioned by inherited and acquired wants of the members of the family, by social customs and standards, and by the array of goods and services from which the family may choose—to mention only the more important of the conditioning influences. Laws of expenditure applicable in one culture and in one epoch may not be applicable in another culture or in another epoch.

---

* This paper was prepared for presentation at the 104th Annual Meeting of the American Statistical Association which was to have been held at Cleveland, Ohio, December 29 to 31, 1942.

Possibly certain broad laws of expenditure can be established as applicable in widely different cultures and epochs, but the present inquiry does not aspire to such very broad generalizations. Its objective is to ascertain what uniformities can be discovered in tendencies of expenditure among families of different sizes, different incomes, and different occupational classes in different regions[1] and in different types of community in the United States in the middle 1930's. The categories of families studied had widely different degrees of education and cultural training. They moved in groups of quite different social standards, and they lived under substantially different climatic conditions. On the other hand, they all belonged, in a broad sense, to a common culture; and they all had available to them substantially the same goods and services at about the same relative prices.

In the final paragraphs of the paper we shall give attention to some evidence on temporal changes in the distribution of family expenditures in the United States. The whole body of evidence suggests a rather remarkable degree of stability in the distribution pattern of family expenditures.

Most persons who have provided data for a questionnaire on family expenditures or who have given thought to the difficulties they would encounter in providing such data entertain some degree of skepticism regarding the accuracy of statistics of expenditures obtained by the survey method. When working with averages from a great number of reports, the investigator need not be concerned over the accuracy of individual reports so long as there is good reason to believe that errors in one direction tend to be about as great, on the average, as errors in the opposite direction. If there exist biassed errors, however, arising from tendencies of reporting families to overstate or to understate their expenditures in certain categories, the effect can be disturbing. Among the sources of error that might exist in the reports, those that would be most serious from the standpoint of the present study would be tendencies to bias that differed among different categories of the population. These would tend to result in evidence that expenditure tendencies differed where in fact they were similar. It would be only by an improbable chance that differing bias in reports made by different groups would result in indicating a common expenditure tendency where the tendencies were in fact substantially different.

The data are found in fact to exhibit a high degree of uniformity in tendencies in the distribution of family expenditure. This finding bears strong testimony to the essential reliability of the data. It does not exclude the possibility, however, that there may exist some bias that appears in about equal degree in the reports by all classes of families. For example, it might be that expenditures in the miscellaneous category tend to be uniformly understated, owing to failure to remember all the small items that should be included. Any concern over possible errors of this sort may well be laid aside until tentative conclusions have been formulated from the data. Few of the conclusions here drawn would be affected to an important degree by any probable bias in the reports.

---

[1] The regional aspect does not enter the present paper explicitly. It is considered in the larger study on which this paper is based and the resulting findings afford part of the basis for the conclusions here stated.

I—GENERAL LAWS

The most general observations that may be made regarding the distribution of family expenditures concern the relation between size of the total expenditure per person and its distribution. Expressing the relationship in this form takes account of the principal effects of three more or less independent influences bearing on the distribution of expenditures: (1) income; (2) size of family; and (3) propensity to save. The proportions of total expenditure devoted to different purposes tend to be about the same for families of the same total expenditure per person even though the families differ with respect to income, size, and proportion of income saved. Certain exceptions to this generalization will be noted below.

As total expenditure per person increases, the proportion of expenditure devoted to food decreases rapidly. At the same time, the proportion of expenditure devoted to recreation and transportation increases greatly, the proportion devoted to clothing increases moderately, and the proportion devoted to a miscellaneous category designated as "other items" increases slightly. The response of expenditure on home maintenance to changing total expenditure per person differs among different categories of the population.

The relationship of expenditure for food to total expenditure is remarkable for the uniformity and closeness with which it approximates the relation

$$F/T = a - b \log T,$$

where F represents expenditure for food and T represents total expenditure. Expressed otherwise, this means that the proportion of total expenditure that is devoted to food tends to decrease exactly in arithmetic progression as total expenditure increases in geometric progression. With only certain exceptions at very low levels of total expenditure, this law of relationship applies closely for families of every size, of every occupation, and in each type of community studied.

One other quite general observation is suggested by the data. Expenditures on recreation and transportation, in the United States in the middle 1930's, tended to reach an upper limit at about one-sixth of total expenditure. Except for low-income farm families, for whom transportation is less appropriately classed with recreation, expenditures on recreation and transportation averaged only about 4 per cent of the total at the lowest levels of total expenditure per person. As total expenditure increased in geometric progression, the proportion devoted to recreation and transportation tended, for all classes of families studied, to increase roughly in arithmetic progression until it reached about 16 per cent of the total, and thereafter to increase but little more.

II—THE CATEGORIES OF EXPENDITURE

The last of the apparent general laws of expenditure stated above required expression in quantitative terms which are dependent to some extent on the precise definition given the category of consumption under discussion. In subsequent consideration of the subject, quantitative expressions will play a large part. The expenditure shown for any category of consumption depends on the definition of the category which is to some extent arbitrary. We may well pause at this point to examine the definitions of the several categories of consumption here used.

For the purposes of the present paper, all expenditures for consumption are distributed among five categories: (1) food; (2) home maintenance; (3) clothing; (4) recreation and transportation; and (5) "other items." The proportion of total expenditure in each of these categories varies greatly with income per person in the family, and varies more or less with other circumstances; but it is useful to observe that, for the "average" family, each of the first two categories—food and home maintenance—comprised about one-third of its total expenditure, and the remaining three categories accounted for about equal parts of the remaining one-third of the total.

Regarding the categories of food and clothing, it need be said only that the value of home-grown food consumed is included, that "food" includes also all beverages, and that clothing expenditures include expenditures for cleaning of clothes, but not for laundry. Under "home maintenance" I count expenditures for housing and for furniture and equipment, and also expenditures for household operation, including payments for household help. This is a category covering the provision of satisfactions much broader than are suggested by the term "shelter" sometimes applied to it.[2] It covers provision for the preparation and serving of food in the home and for the laundry of clothing, and it covers much of the cost of providing those amenities of life that are essentially recreation enjoyed at home. It might be desirable to have a category of expenditures for shelter only, if that could be done; but the narrowest available category—that for "housing"—covers a great deal beyond expenditures for mere shelter, and covers more under some circumstances of occupancy of the family dwelling than under other circumstances. Under these conditions I have chosen to broaden the category to comprehend substantially all those expenditures that contribute toward affording the various satisfactions associated with housing.

Recreational expenditures are defined in the data here used to include among other things: expenditures for admissions of family and guests to movies, lectures, ball games or dances; expenditures for participation in sports and games (but not including sports clothing); expenditures on radio, music, photography and pets; dues to social and recreational clubs.[3] With these I include the much larger expenditures on automobile and other transportation. Certain minimum transportation expenses, varying in amount from family to family, are quite unrelated to recreation, but expenditures for transportation above such minima may be regarded as essentially recreational expenditures.

In the miscellaneous category, "other items," the largest single component, accounting for about half of the total at each expenditure level, is medical care, including drugs. Other components are tobacco, toilet supplies, cosmetics, and personal services at barber shops and beauty parlors; education and special lessons in sports, music and dancing; books, newspapers and magazines; and some expenses not suggested by this list. Gifts and taxes are not here counted as expenditures for consumption.

III—CONDITIONS OF LIFE AND SOCIAL STANDARDS IN RELATION TO EXPENDITURE PATTERNS

Certain of the conditions under which a family lives and works influence its distribution of expenditures, and so also do the standards of the social group to which it belongs. The effects of these two types of influence are not in general clearly separable on a statistical basis, since differences in living conditions as actually observed are generally associated with differences in social standards also, and there are lacking quantitative measures of living conditions or of social standards such as might permit use of multiple correlation analysis to ascribe to each its special influence. The effects of differences in living conditions and in social standards are readily observable in differences of expenditure patterns between urban, village and farm families, and in differences according to occupation. To some extent occupation is itself an index of living conditions. A special type of difference in living conditions is reflected in size of family.

---

[2] As in the publications of the National Resources Planning Board, for example, I use the term "home maintenance" following the practice of the Bureau of Labor Statistics. In the case of families owning their homes, estimated rental value is included.

[3] See National Resources Planning Board, *Family Expenditures in the United States: Statistical Tables and Appendixes*, 1941, p. 143.

Differences in conditions of life and in social standards, so far as they are reflected in the bases of classification that I have employed, appear to have exerted their chief direct influence on expenditure for home maintenance. They were necessarily accompanied by opposite differences in percentages of expenditure devoted to other categories of consumption. In some instances differences in expenditure for home maintenance were offset largely by differences in expenditure for food, and in other instances mainly or entirely by differences in expenditures on recreation and transportation and on clothing. It appears also that comparisons between types of community must take into account differences in general cost of living that give the dollar a greater value in rural districts than it has in large cities.

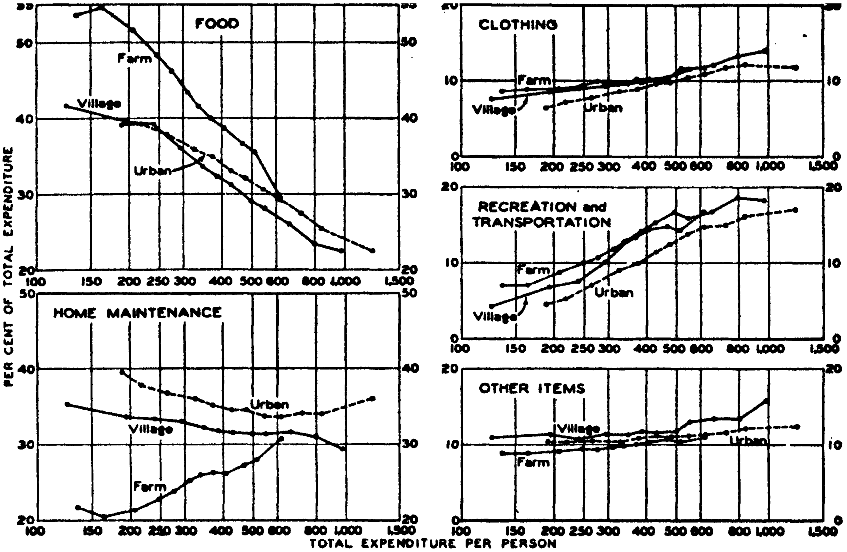
Effect of type of community.—As regards the proportions of family expenditure devoted to the three minor categories of consumption, extraordinarily little difference appears between urban, village, and farm families. The data are shown graphically in the right half of Chart I in terms of curves which reflect the relation of total expenditure per person to the per cent of the total falling into individual categories. In each type of community the relations accord with the general laws stated in Section I.

Urban families spent less on recreation and transportation and less on clothing at each level of total expenditure per person than did either village or farm families. This rather surprising fact appears in a more reasonable light when it is noted that the percentages spent by urban families for recreation and transportation and for clothing at each level of total expenditure per person tended to be about the same as the percentages spent for the same purposes by village and farm families with a total expenditure per person 20–25 per cent smaller. The general cost of living is substantially lower in villages and in rural communities, as a rule, than in large cities, and it may well be that if comparison could be made between families with the same purchasing-power expenditure per person rather than between families with the same dollar expenditure per person, the percentages spent by urban, village and farm families for recreation and transportation and for clothing would be found to have differed very little.

CHART I

EXPENDITURE CURVES FOR URBAN, VILLAGE, AND FARM FAMILIES*

(*Arithmetic scale for percentages; logarithmic scale for total expenditure per person.*)



* Data computed from tables in *Family Expenditures in the United States: Statistical Tables and Appendixes*, National Resources Planning Board, 1941.

The close similarities among urban, village and farm families as regards tendencies in allocation of expenditure to recreation and transportation, clothing and "other items" make it appear the more remarkable that they differ so greatly as regards expenditures for home maintenance and for food. Among urban and village families spending a total of about $200 per person, slightly less than 40 per cent of the total was spent for food. For farm families at the same expenditure level the food consumed was valued at over 50 per cent of the total expenditure.[4]

[4] The imputed value of home-grown food is included in the figures for total expenditure as well as in the figures for expenditure for food.

On the other hand, the proportions of total expenditure that were devoted to home maintenance by families spending about $200 per person were about 39 per cent for urban families, 33 per cent for village families, and only 21 per cent for farm families. If the data from which these percentages were derived represented only a few families in a few selected communities, these differences might deserve little attention. The percentages are based, however, on records for a great number of families in numerous communities widely scattered over the United States. The smooth course of the corresponding curves in the left half of Chart I attests the consistency of the data. The disparities among the curves unquestionably reflect existence of striking differences in expenditure tendencies among urban, village and farm families in the United States in the middle 1930's.

The differences between urban families and farm families as regards expenditures on food and on home maintenance are complementary, and both may arise mainly from conditions that encourage low expenditure on living accommodations on the farm. A minor factor in the difference is land rent. The ground on which the farm home stands is cheap by comparison with the price of a city lot. Indeed, land value may rather generally have been neglected in estimating the cost of home maintenance on farms. There is a tendency also for construction costs to be lower in the country than in the city and for old farmhouses to be valued much farther below replacement cost than similarly old dwellings in cities. There remains, however, the major fact: that farm families at a low economic level tend to be content with poorer living quarters, less well furnished, than city families at the same economic level. In a very real sense the farm itself is a home. On a farm, the family whose dwelling affords little but the minimum of shelter enjoys pleasures of living comparable with those to be obtained in a city only through substantially larger expenditure on the home itself.

The saving which farm families tend to make on expenditure for home maintenance is not distributed in equal proportions among other categories of consumption, but tends to be reflected mainly in increased outlay for food. The large energy requirements of most farm work lead to heavy food consumption, home-produced foods tend to be used generously, and table waste, because it provides food for the farm animals, is not watched as closely as with most city families in similar circumstances.

As total expenditure increases, the expenditure pattern of farm families tends to approach that of city families. At an expenditure level of

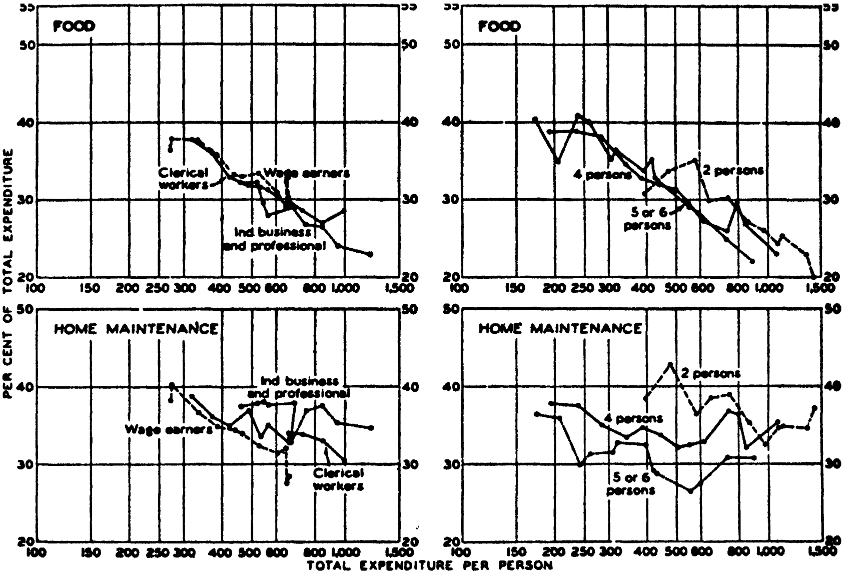
$600 per person during the middle 1930's, the distribution of expenditures by farm families differed little from that of city families.

Expenditures of village families on food and home maintenance show characteristics much closer to urban families than to farm families. The differences between village and urban families as regards expenditures on food and on home maintenance may arise mainly from differences in costs.

CHART II

FOOD AND HOME MAINTENANCE EXPENDITURE CURVES FOR COLUMBUS, OHIO, FAMILIES CLASSIFIED BY OCCUPATION AND BY SIZE*

(*Arithmetic scale for percentages; logarithmic scale for total expenditure per person.*)



* Data computed from tables in *Family Income and Expenditure in Nine Cities of the East Central Region, 1935–36: Vol. II, Family Expenditure* (Study of Consumer Purchases: Urban Series, U. S. Bureau of Labor Statistics Bull. 644, 1941.)

The tendency for the proportion of expenditure devoted to food to decrease in arithmetic progression as total expenditure increases in geometric progression appears clearly in Chart I. With percentages plotted to an arithmetic scale and total expenditures to a logarithmic scale, all three curves of expenditure for food approximate closely to straight lines for total expenditures above $230 or less per person.

Effect of occupation.—The special expenditure characteristics of farm families are perhaps attributable quite as much to occupation as to place of residence. Among urban families, occupation had a considerable influence on the distribution of expenditures, especially as regards expenditure on home maintenance. In the left half of Chart II are shown graphically the relations between total expenditure per person and the proportions of the total devoted to food and to home maintenance by Columbus, Ohio, families in three occupational classes. The expenditure characteristics noted below in discussion applying specifically to the data shown for Columbus have been found also in similar data for New York City, though with differences as regards the actual percentages at particular expenditure levels.[5]

Independent business and professional families tended to devote about 37 per cent of total expenditure to home maintenance at most expenditure levels, but there appears to have been a slight tendency for the proportion to decline as total expenditure increased, from nearly 38 per cent for families spending under $700 per person to 35 per cent for families spending about $1,000 per person.

Wage earners and clerical workers at the lowest levels of expenditure per person devoted over 38 per cent of their expenditures to home maintenance. The proportions declined markedly and in approximate arithmetic progression as total expenditure increased in geometric progression. At the highest levels of expenditure per person recorded for families of wage earners and clerical workers, they devoted about 30 per cent or less to home maintenance. At each level of total expenditure, the proportion devoted to home maintenance by clerical workers was above that for wage earners and below that for independent business and professional families.

The differences between occupations as regards expenditure on home maintenance are not reflected in opposite differences in expenditure on food, the other major category of expenditure (Chart II). Instead, they were counterbalanced almost entirely by opposite differences in expenditures on recreation and transportation (data not shown). Much expenditure on home maintenance may be essentially expenditure for recreation. Thus the data suggest the inference that business and professional families tend to take more of their recreation in the home than do families of wage earners and clerical workers with the same total expenditure per person. It is to be noted, however, that among farmers—another occupational category—low expenditure on home maintenance does not go with high expenditure on recreation in the formal sense.

---

[5] The statistics have been so published that data such as are shown in Chart II can be obtained only by calculation for each city separately. In the published tables, the classification is by family income rather than by expenditure per person. The lines between plotted points in Chart II connect points corresponding to successive classes in the tabulation by family income.

Effect of size of family.—Family size shows effects on the distribution of expenditure that arise partly from inherent "advantages of size." The large family can make more efficient use of normal housing facilities and of the appurtenances of the home than can the small family. This seems the principal explanation of the large differences in expenditure on home maintenance shown by the curves of Chart II for families of different size. A tabular comparison for Columbus families having a total expenditure of $550–$650 per person appears as follows:

Size and description of family	Proportion for home maintenance
2, husband and wife	40 per cent
4, husband, wife, and 2 children under 16	32 per cent
5–6, husband, wife, 1 child under 16, 1 person over 16, and 1 or 2 others of any age	27 per cent

For each size of family, the data show a decline in proportion of expenditure for home maintenance as total expenditure increased up to a certain point, and then a tendency for the proportion spent on home maintenance to increase slightly with further increase in total expenditure. The combined figures for all urban families as shown in Chart I exhibit even more clearly this characteristic of expenditure on home maintenance. It results principally at least from the fact that at the lower and middle expenditure levels the percentages are dominated by data for families of wage earners and clerical workers, while at the upper expenditure levels the percentages are dominated by data for families of business and professional people. On this account it cannot be said that there exists a general tendency for the curves of percentage expenditure on home maintenance to be concave upward. Rather, there are different tendencies for different occupations, and for each occupation the expenditure curve, as previously noted, is nearly linear when plotted as in the accompanying charts.

It is noteworthy also that for families spending over about $800 per person there was little or no difference between percentages spent on home maintenance by two-person and by four-person families. Although larger families use the facilities of the home more efficiently than two-person families, they have also more inclination to expand the satisfactions that the home affords. As between two-person and four-person families, these opposing influences offset each other at the higher levels of expenditure per person.

Expenditures for food by Columbus families of different sizes show a tendency for two-person families to spend more on food than did larger families with the same expenditure per person. The reason is not clear. Two-person families may tend to waste more food, they may tend to entertain more at meals, and/or they may tend to use more expensive food, bought in cans or at delicatessen stores. The logical suggestion that they may tend to spend more on purchased meals is contradicted by the statistics.[6]

The curves of expenditure for food by families of different size show again the characteristic form of such curves. The proportion spent for food tends to decline in arithmetic progression as total expenditure increases in geometric progression.

The relatively high expenditures by two-person families on both home maintenance and food were made possible in about equal degree by holding down expenditures on recreation and transportation and on clothing. Families of five or six persons with total expenditures under $400 per person took advantage of their economy on home maintenance chiefly to increase expenditure on clothing, but at higher levels of total expenditure they also divided the difference about equally between increased expenditure on clothing and increased expenditure on recreation and transportation.

IV—TEMPORAL CHANGES IN TENDENCIES OF FAMILY EXPENDITURE

The foregoing observations may be used in conjunction with calculations provided by W. Allen Wallis in a recent paper[7] to throw some light on changes that occurred in tendencies of family expenditure in the United States during the first 35 years of the present century. For this purpose we may use most advantageously Wallis' data on the amounts expended for certain categories of consumption by families at an average income level.[8]

Such great changes in value of the dollar as have occurred since the beginning of the century may be assumed to tend to cause horizontal shifts in such expenditure curves as are shown in Charts I and II above. Such shifts would tend to be accompanied by large changes in the percentage distribution of expenditures at any particular expenditure or income level. In interpreting Wallis' data, this circumstance must be taken into account. For the present, I attempt only such conclusions as may be drawn without making quantitative adjustment for the price-level factor.

---

[6] *Family Income and Expenditure in Nine Cities of the East Central Region, 1935–36: Vol. II, Family Expenditure* (Study of Consumer Purchases: Urban Series, U. S. Bureau of Labor Statistics Bull. 644, 1941), pp. 133–135.

[7] "The Temporal Stability of Consumption Patterns," *The Review of Economic Statistics*, November 1942, XXIV, 177–183.

[8] See his figures in Column H, Table I, of the paper cited above.

A striking feature of the data presented by Wallis is the great increase between 1901 and 1935 in the principal components of expenditure for home maintenance. In terms of percentages of income reported, expenditure by urban families for rent or its equivalent and for fuel and light increased from 18.9 per cent in 1901 to 29.1 per cent in 1935. Expenditure for "household operation" increased between 1918 and 1935 from 8.7 per cent to 12.1 per cent. These increases seem greater than can be explained by such horizontal shifts in the expenditure curves as might be expected to accompany the changes in general cost of living that occurred over the interval.

Expenditures for recreation and transportation increased from 6.3 per cent in 1918 (they are not given for earlier years) to 13.7 per cent in 1935. This seems also a considerably larger increase than can be accounted for by a shift in expenditure curves resulting only from change in the cost of living.

It appears then that, over the period for which we have data on family expenditures in the United States, there have been some marked changes in tendencies of family expenditure. If the proportions of expenditure devoted to home maintenance and to recreation and transportation have tended to increase, it is likely that the proportions devoted to food and to clothing have tended to decrease. Changes in relative costs have doubtless exercised considerable influence on the result. Detailed interpretation must wait on solution of some difficult problems of quantitative comparison of expenditure patterns at different times.

V—APPLICATIONS OF STATISTICAL LAWS OF FAMILY EXPENDITURE

Section I above is essentially a summary of the main conclusions developed in the sections that followed. At this point it is appropriate to look at the significance of the findings from a different point of view. To what extent are the results of surveys of family expenditures applicable to the interpretation of current economic situations and to the anticipation of future developments?

Tendencies in family expenditure do change with time and no studies yet made provide a reliable basis for appraising current changes or anticipating future changes. Until the requisite studies are made, changes must be allowed for on the basis of *a priori* judgments, unreliable as they may be. To such judgments the findings of the present paper may contribute appreciably.

It is particularly noteworthy in this connection that the present findings support the view that the effects of changes in income through time may be judged fairly reliably from differences in expenditure patterns at different levels of expenditure at a given time. This is indicated especially by the fact that expenditure patterns differ little according to size of family, and that such differences as are observed are explicable largely on technical grounds. Expenditure per person depends to a large extent on size of family, which changes fairly rapidly through time for individual families. The expenditure patterns for families of different sizes might show differences resulting from recent reduction in expenditure per person incident on increase in the family size. The fact that differences calling for such explanation do not appear suggests that families allowed or compelled to change to a new level of expenditure per person tend to change quickly to the expenditure pattern characteristic of that level.

Changes in income through time probably tend to have a pronounced effect on the proportion of income saved. Curves expressing the relation of expenditures in particular categories to total expenditure per person, therefore, are probably more stable in periods of changing income than are curves expressing the relation of expenditures in particular categories to income per person.

CORRECTION FORMULAS FOR MOMENTS OF A GROUPED-DISTRIBUTION OF DISCRETE VARIATES

By Joseph A. Pierce
Atlanta University

Because of the labor involved in the computation of moments of certain frequency distributions of discrete variates, it is often desirable to consider a grouped-distribution which is obtained from the original distribution by an arbitrary grouping of the given data in class intervals of width k. The moments are then computed on the assumption that all of the items of each class are concentrated at the middle point of the interval. This assumption introduces grouping errors because the items of each interval usually do not all lie at the middle point of the interval. It is necessary, therefore, that corrections be used which compensate for the grouping errors.

The general formula for correction for grouping errors of a distribution of discrete variates which has been grouped in "k groupings of k" has been given by the writer[1] as

$$\mu_{1:\mu_t} = \sum_{s=0}^{[t/2]} \binom{t}{2s} \mu_{t-2s} R_{2s}, \tag{1}$$

where $\mu_{1:\mu_t}$ is the average of the t-th moments of the grouped-distributions, μ_{t-2s} is the $(t-2s)$-th moment of the ungrouped data and R_{2s} is the $2s$-th central moment of a rectangular distribution of discrete variates. Usually it is desired to obtain the value of the moments of the ungrouped data but (1) solved explicitly for μ_t takes a complicated form[2] and is omitted from this paper. However, it may be readily seen that these correction formulas express the true moments in terms of the average grouped moments. Thus the corrections which are found are *average corrections*. But in actual problems we do not know the exact values of the averages of moments of the grouped-distributions and we approximate their values by utilizing the corresponding moments computed from a single, arbitrarily selected grouping. It is possible, in any instance, that the use of the approximation may cause a greater deviation between the "corrected" moment and its true value than would have resulted if no attempt were made to correct for grouping. Thus, while the systematic errors due to grouping are removed by the correction formulas, the problem of eliminating the accidental errors, introduced through the use of approximations for the averages of moments of grouped-distributions, has not been solved. Carver[3] pointed out the serious nature of this problem and recently Dwyer[4] made an advance towards its solution by presenting grouping methods which give better "corrected" values of moments from a single grouping than can be obtained with the use of the conventional correction formulas.

---

[1] J. A. Pierce, "A Study of a Universe of n Finite Populations with Application to Moment-Function Adjustments for Grouped Data," *Annals of Mathematical Statistics*, Vol. 11 (1940), p. 326.

[2] *Op. cit.*, p. 327.

It is the purpose of this paper to present new correction formulas for moments, computed from a single grouping, which yield the exact values of the moments of the ungrouped data.

THE FIRST MOMENT

Let the ungrouped distribution be the set of variates x_i, $(i=1, 2, \cdots, s)$, with the corresponding frequencies f_i, $(i=1, 2, \cdots, s)$. The x's are subject to the condition $x_{i+1}-x_i=h$ and consequently some of the f's may be zero. The mean of this distribution is $\mu_1=\sum xf/N$ where $N=\sum f$.

If the given distribution is grouped in a "grouping of k" and any x_i is placed in the r-th position of a class, then the limits of the class are $x_i-(r-1)h$ and $x_i+(k-r)h$ and the midvalue is $x_i+\frac{1}{2}[k-(2r-1)]h$. Therefore, when the midvalue is used as the value of x_i, the quantity $\frac{1}{2}[k-(2r-1)]h$ is added to the true value of x_i. For convenience, we define

$$e_r = \tfrac{1}{2}[k-(2r-1)]h \tag{2}$$

and

$$F_r = f_r + f_{r+k} + f_{r+2k} + \cdots. \tag{3}$$

If v_i is the midvalue of the i-th class in the r-th grouped-distribution, the mean of the ungrouped data is

$$\mu_1 = \frac{1}{N}\sum_{l=0}^{k-r}(v_1 - e_{r+l})f_{l+1} + \frac{1}{N}\sum_{i=2}^{c}\sum_{j=(i-1)k+1}^{ik}(v_i - e_j)f_{j-r+1} \tag{4}$$

where c is the number of classes and $e_{k+p}=e_p$, $p=0, 1, 2, \cdots, (r-1)$. Therefore

$$\mu_1 = {}_r\mu_1 - \frac{1}{N}\sum_{i=1}^{k}F_i e_{r+i-1}, \quad (r = 1, 2, \cdots, k), \tag{5}$$

or

$$\mu_1 = {}_r\mu_1 - \phi_r^{(1)}, \quad (r = 1, 2, \cdots, k), \tag{6}$$

---

[3] H. C. Carver, "The Fundamental Nature and Proof of Sheppard's Adjustments," *Annals of Mathematical Statistics*, Vol. 7 (1936), p. 171.

[4] P. S. Dwyer, "Grouping Methods," *Annals of Mathematical Statistics*, Vol. 13 (1942), pp. 138–155.

where we define

$$\phi_r^{(n)} = \frac{1}{N} \sum_{i=1}^{k} F_i e^n_{r+i-1} \tag{7}$$

and ${}_r\mu_1$ is the first moment of the grouped-distribution in which x_1 is in the r-th position of a class.

It is obvious that we may form $\phi_{r+1}^{(n)}$ from $\phi_r^{(n)}$ by a cyclic permutation of the e's of $\phi_r^{(n)}$.

If $F_1 = F_2 = F_3 = \cdots = F_k$, it follows that $\phi_r^{(1)} = 0$ for all r. Therefore, if k can be selected such that all of the F's are equal, then the mean of any grouped-distribution equals the mean of the ungrouped data.

Also from (6) we are able to find the difference between the mean of any grouped-distribution and the mean of the ungrouped data without computing either mean. Therefore, if one desires to present the data in the grouping for which the mean differs least from the true mean then it is only necessary to select the grouping which corresponds to the smallest $\phi_r^{(1)}$.

HIGHER MOMENTS

The t-th moment of the ungrouped distribution is

$$\mu_t = \frac{1}{N} \sum_{l=0}^{k-r} (v_1 - e_{r+l})^t f_{l+1} + \frac{1}{N} \sum_{i=2}^{c} \sum_{j=(i-1)k+1}^{ik} (v_i - e_j)^t f_{j-r+1} \tag{8}$$

which after some simplification, becomes

$$\mu_t = (-1)^{t-1} {}_r\mu_t - \frac{1}{N} \sum_{j=1}^{t-1} (-1)^{t+j} \binom{t}{j} \sum^{i} v_i^{t-j} V_i^{(j)} + (-1)^t \phi_r^{(t)} \tag{9}$$

where $V_i^{(j)}$ is the sum of the j-th power of the variates which were grouped in the i-th class and $\sum^{i}$ operates on the classes of the grouped-distribution.

ILLUSTRATION

We use the arbitrary distribution of Table I to illustrate the method of finding the moments of an ungrouped distribution from the moments of a grouped-distribution. We arbitrarily choose a class interval of three and hence the data of Table I are written in Table II in columns of three rows. The F's, as defined by (3), are computed by simply adding the f's of each row. Thus $F_1 = 33$, $F_2 = 13$, and $F_3 = 14$.

Since terms of the form $\sum v_i^t V_i^{(j)}$ appear in the correction formulas, where $V_i^{(j)}$ represents the sum of the j-th power of the variates which fall in the i-th class of the chosen grouped-distribution, we find it necessary to include columns of $V_i^{(j)}$ in the grouped-distribution table.

TABLE I

Variate x	Frequency
1	2
2	8
3	10
4	30
5	4
6	3
7	1
8	1
9	1

TABLE II

x	f	x	f	x	f	F
1	2	4	30	7	1	2+30+1=33
2	8	5	4	8	1	8+4+1=13
3	10	6	3	9	1	10+3+1=14

Jones[5] and Dwyer[6] suggest that a column of $V_i^{(1)}$ be included but for accuracy in the t-th moment it is necessary to use $V_i^{(t-1)}$.

However, it is usually not necessary to find moments higher than the fourth and in many instances the first and second moments are sufficient.

Although it is possible to form three groupings of three, in practice one obtains the moments from a single grouping. Formula (9) applies equally well to any grouping. Therefore, there is no loss of generality when we select the grouping in which the first variate is in the first position of the class. This grouped-distribution, together with $V_i^{(j)}$, $j=1, 2, 3$, is given in Table III.

TABLE III

Class	Midvalue v	f	$V^{(1)}$	$V^{(2)}$	$V^{(3)}$
1–3	2	20	48	124	336
4–6	5	37	158	688	3068
7–9	8	3	24	194	1584

The correction formulas through the fourth moment, as given by (9), are

$$\mu_1 = {}_r\mu_1 - \phi_r^{(1)}, \tag{10}$$

$$\mu_2 = -\,{}_r\mu_2 + \frac{2}{N}\sum v_i V_i^{(1)} + \phi_r^{(2)}, \tag{11}$$

$$\mu_3 = {}_r\mu_3 - \frac{3}{N}\sum v_i^2 V_i^{(1)} + \frac{3}{N}\sum v_i V_i^{(2)} - \phi_r^{(3)}, \tag{12}$$

$$\mu_4 = -\,{}_r\mu_4 + \frac{4}{N}\sum v_i^3 V_i^{(1)} - \frac{6}{N}\sum v_i^2 V_i^{(2)} + \frac{4}{N}\sum v_i V_i^{(3)} + \phi_r^{(4)}. \tag{13}$$

---

[5] H. L. Jones, "The Use of Grouped Measurements," this JOURNAL, Vol. 36 (1941), p. 529.

[6] P. S. Dwyer, *ibid.*, p. 154.

Before we can apply the correction formulas we must compute $\phi_1^{(1)}$, $\phi_1^{(2)}$, $\phi_1^{(3)}$, and $\phi_1^{(4)}$. We get

$$\phi_1^{(1)} = \frac{1}{N}\,[e_1F_1 + e_2F_2 + e_3F_3] = \frac{19}{60},$$

$$\phi_1^{(2)} = \frac{1}{N}\,[e_1^2F_1 + e_2^2F_2 + e_3^2F_3] = \frac{47}{60},$$

$$\phi_1^{(3)} = \frac{1}{N}\,[e_1^3F_1 + e_2^3F_2 + e_3^3F_3] = \frac{19}{60},$$

$$\phi_1^{(4)} = \frac{1}{N}\,[e_1^4F_1 + e_2^4F_2 + e_3^4F_3] = \frac{47}{60}\,.$$

We now substitute in the correction formulas and find $\mu_1 = \frac{230}{60}$, $\mu_2 = \frac{1006}{60}$, $\mu_3 = \frac{4988}{60}$, $\mu_4 = \frac{28066}{60}$, which are identical with the moments of the ungrouped data.

Comparisons between moments obtained by various methods may be made from Table IV.

TABLE IV

Dist. \ Moments	μ_1	μ_2	μ_3	μ_4
Ungrouped dist.	$\frac{230}{60}$	$\frac{1006}{60}$	$\frac{4988}{60}$	$\frac{28066}{60}$
Grouped dist., uncorrected	$\frac{249}{60}$	$\frac{1197}{60}$	$\frac{6321}{60}$	$\frac{35733}{60}$
Grouped dist., corrected by usual formulas	$\frac{249}{60}$	$\frac{1057}{60}$	$\frac{5823}{60}$	$\frac{31465}{60}$
Grouped dist., corrected by new formulas	$\frac{230}{60}$	$\frac{1006}{60}$	$\frac{4988}{60}$	$\frac{28066}{60}$

LIMITS OF ERROR BETWEEN GROUPED MOMENTS AND UNGROUPED MOMENTS

If the limits of error between the desired moment of the ungrouped data and the corresponding moment for any grouped-distribution are close, it may not be worth while to use the exact formulas of this paper.

Norton[7] gives formulas for finding the extreme errors in the moments of a continuous distribution. Formulas for the maximum difference between μ_t and ${}_r\mu_t$ for a distribution of discrete variates are derived below.

---

[7] K. A. Norton, "Limits to the Accuracy of Estimated Moment Coefficients," *Sankhya*, Vol. 3 (1938), pp. 265–272.

The value of μ_t, denoted by μ_t^+, for which $\mu_t - {}_r\mu_t$ is a positive maximum occurs when the difference between v_i and each variate in the class is the largest possible value, namely e_1, and positive. In other words, $\mu_t - {}_r\mu_t$ is a positive maximum when the variates of the classes are all at the upper limits of the classes. Therefore,

$$\mu_t^+ \quad \frac{\sum (v_i + e_1)^t f_i}{\sum f_i}.$$

Similarly, the value of μ_t, denoted by μ_t^-, for which $\mu_t - {}_r\mu_t$ is a negative maximum is given by

$$\mu_t^- = \frac{\sum (v_i - e_1)^t f_i}{\sum f_i}$$

and necessarily

$$\mu_t^- \leqq \mu_t \leqq \mu_t^+.$$

Therefore

$$\sum_{i=0}^{t} (-1)^i \binom{t}{i} {}_r\mu_{t-i} e_1^i \leqq \mu_t \leqq \sum_{i=0}^{t} \binom{t}{i} {}_r\mu_{t-i} e_1^i \tag{14}$$

which, for the first two moments, becomes

$${}_r\mu_1 - e_1 \leqq \mu_1 \leqq {}_r\mu_1 + e_1, \tag{15}$$

$${}_r\mu_2 - 2{}_r\mu_1 e_1 + e_1^2 \leqq \mu_2 \leqq {}_r\mu_2 + 2\mu_1 e_1 + e_1^2. \tag{16}$$

It follows that (14) may be used to determine the maximum deviation of a given moment of the grouped-distribution from the corresponding moment of the ungrouped data.

CONCLUSION

If precision is not desired or if the limits of error between ${}_r\mu_t$ and μ_t are close, the conventional correction formulas (1) may prove satisfactory. But if additional precision is needed, the new correction formulas of this paper should be used. In order that these new formulas may be used it is not sufficient to record the results of grouping in the form of classes and frequencies The grouped-distribution should also contain $F_1, F_2, \cdots, F_k$ and columns of $V_i^{(n)}$. Usually it will not be necessary to go beyond $V_i^{(1)}$. This additional information requires only a little more space and labor yet it makes it possible for others who compute moments from the grouped-distribution to do so confident that their results are accurate. It is to be hoped that this procedure will be followed by those who prepare grouped-distributions for general use.

A SECOND MOMENT CORRECTION FOR GROUPING

By G. R. Davies and Nancy Bruner
University of Iowa

It is a matter of practical experience that the use of Sheppard's correction in estimating the population variance of grouped sample data sometimes results in an unreasonably low figure, particularly when the sample distribution lacks "high contact," that is, when the first and last frequencies are relatively large. This downward bias obviously arises from the fact that the correction theoretically assumes and corrects for classes from minus infinity to plus infinity, as in the case of the normal curve.

Several attempts have been made to approximate and remove this bias, but none of the formulas proposed has come into general use.[1] This failure arises in part from the difficulty encountered in devising a smoothing procedure which is at once as plausible as Sheppard's, and at the same time is convenient to use. An even greater difficulty is that of adapting the correction to a discrete scale having smaller class intervals than those of the given distribution, as is desirable when a measure like "persons per dwelling," grouped to include a range of several persons, is to be reduced to discrete unit classes.

It is the purpose of this paper to develop a correction which will be identical with Sheppard's at the limiting case of a continuous scale and high contact, and yet will be adaptable either to an integral number of equally wide sub-classes, or to a continuous scale. At the same time the correction must not be unduly influenced by irregularities in the data, outside of the first and last frequencies, which necessarily measure lack of high contact.

It will be found convenient to approach the problem by observing that Sheppard's second moment correction in effect compensates for two distinct errors arising from grouping. The first of these two errors is the result of considering the squared class mark as representative of the average of corresponding squares in the subclasses. Thus a frequency of f, with class limits of 0 and 8, is usually assumed to contribute to the calculation of the second moment $fm^2 = f \times 4^2 = 16f$. But if circumstances require that it should be broken down into two equal subclasses of one-half the original width, it contributes $f(2^2+6^2)/2 = 20f$, or, similarly, if into four subclasses, $f(1^2+3^2+5^2+7^2)/4 = 21f$.

---

[1] W. T. Lewis in "A Reconsideration of Sheppard's Corrections," *Annals of Mathematical Statistics*' Vol. 6, 1935, has derived a complex second moment correction somewhat similar in principle to that here developed. Also see the September, 1940, issue, pp. 324–332.

The sum of squares, corrected for the grouping error thus illustrated, is[2]

$$fm^2_{cs} = fm^2 + f\cdot\frac{i^2}{12}\cdot\frac{k^2-1}{k^2} \tag{1}$$

where f is the frequency, m is the class mark, i is the class interval, k is the integral number of subclasses in one class, and m_{cs} is the class mark corrected for squares, that is, adjusted for subclasses as just illustrated. It is obvious that if the scale is continuous and k is assumed to be infinite the fraction $(k-1)/k^2$ disappears, and, in the above illustration,

$$fm_{cs} = fm^2 + \frac{fi^2}{12} = 16f + \frac{8^2}{12}f = 21.33f.$$

If the class intervals of a given distribution have the same width throughout, and if each class is broken down into the same number of subclasses, the aggregate correction for the whole distribution, with respect only to the squared class marks, is

$$\sum fm^2_{cs} = \sum fm^2 + N\cdot\frac{i^2}{12}\cdot\frac{k^2-1}{k^2}\,\cdot \tag{2}$$

It may be noted that if k of this formula is again assumed to be infinite, that is, if $(k^2-1)/k^2$ is taken at its limit of unity, the correction term follows Sheppard's second moment correction, except that the sign is positive, suggesting that for continuous, high-contact distributions another correction, twice as great and negative, is required. The nature of this second correction will now be considered.

The second correction for grouping makes allowance for the error involved in the usual assumption that in each class of a distribution the frequencies are distributed equally throughout the class, instead of being somewhat more numerous on the side of the class nearest to the mode. This may be called a correction for slope or height, inasmuch as graphically, on a continuous scale, it changes the flat top of the frequency rectangle into a curve approximating the smoother outline of the whole distribution. It will be seen that this correction is made sepa-

---

[2] This formula is readily obtained as follows: When k is even,

$$\frac{f}{k}\left[\left(m+\frac{i}{2k}\right)^2+\left(m+\frac{3i}{2k}\right)^2+\cdots+\left(m+\frac{i(k-1)}{2k}\right)^2 \right.$$
$$\left. +\left(m-\frac{i}{2k}\right)^2+\left(m-\frac{3i}{2k}\right)^2+\cdots+\left(m-\frac{i(k-1)}{2k}\right)^2\right]$$
$$=\frac{f}{k}\left[km^2+\frac{2i^2}{4k^2}\left(1^2+3^2+5^2+\cdots+(k-1)^2\right)\right]=fm^2+\frac{fi^2(k^2-1)}{12k^2}\,.$$

A similar arrangement for k odd gives the same result.

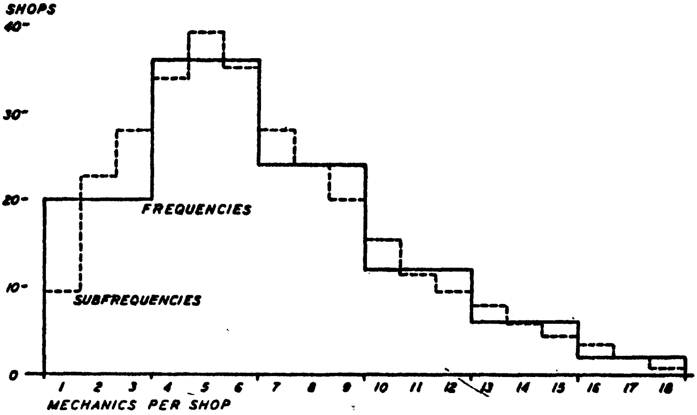
rately in each frequency, which distinguishes the procedure from the more radical one of curve fitting.

Stated more precisely, the second correction applied to a given frequency involves a readjustment of the graphic height of each of the k subfrequencies in such a way that they conform to the outline of the distribution and yet retain in the aggregate the original area of the frequency rectangle. The plotted subfrequencies in the accompanying chart illustrate the nature of the adjustment. This graduation of the subfrequencies may be most conveniently accomplished by making them conform to certain parabolas about to be described, inasmuch as a normal or logarithmic normal curve could be closely approximated by a series of suitable short parabolas. Such parabolas may be constructed, one within each class, to form a connected series beginning and ending with zero height ($Y=0$) at the limits of the distribution, each two adjacent parabolas meeting at the average height of the two frequencies they respectively represent. Each rectangular subclass is then given the same area as that under the portion of the parabola having the same X limits. The height of the parabola may then be adjusted to pass through the central ordinates of the given class subfrequencies at the graphic height of each f_s.

CHART I

CLASSIFICATION OF CERTAIN SHOPS ACCORDING TO NUMBER OF MECHANICS EMPLOYED (SEE TABLE I)

Frequencies reported in classes of three, and reduced to unit classes by parabolic interpolation. Plotted subfrequencies, read on Y scale, should be divided by $k=3$.



Any one of the parabolas thus adjusted may be defined in terms of the given frequency (f) of its own class, the preceding frequency (f_{-1}), and the following frequency (f_{+1}) by the equation

$$Y' = \frac{2f - f_{+1} - f_{-1}}{8} \cdot \frac{k^2 - 1}{k^2} + \frac{f_{+1} - f_{-1}}{2i} x + \frac{3(f_{+1} + f_{-1} - 2f)}{2i^2} x^2 \quad (3)$$

where x has its origin at the class mark of the given frequency, and the "base" ($Y' = 0$) of the parabola is f. If k is infinite, the integral of $Y'dx$ between the class limits $x = -i/2$ and $x = +i/2$ will then equal zero, as required.

Multiplying each squared subclass mark (m_s^2) in a given class by its subfrequency (Y'_s/k), summing and simplifying, give the total,

$$\sum Y_s'm_s^2/k = fm_{ch}^2 = \left(2bm + \frac{ci^2}{15} \cdot \frac{k^2 - 4}{k^2}\right) \frac{i^2}{12} \cdot \frac{k^2 - 1}{k^2} \quad (4)$$

This total embodies the correction for slope (graphic heights of subfrequencies) and for the class-mark squares in so far as they enter into the modification of height. The two corrected total squares, (1) and (4), may now be combined, and the frequency values of b and c, as indicated in (3), substituted. The result is a corrected total of squares for any class except the two at the extremes of the distribution, as follows:

$$fm_c^2 = fm^2 + \left[f - \frac{m}{i}(f_{-1} - f_{+1}) - \frac{2f - f_{-1} - f_{+1}}{10} \cdot \frac{k^2 - 4}{k^2}\right] \frac{i^2}{12} \cdot \frac{k^2 - 1}{k^2} \cdot \quad (5)$$

While a formula for the total corrected squares might be obtained more directly by expanding $\sum Y_s\, m_s^2/k$, where Y_s expresses the full graphic height of the subfrequencies, the method here adopted, involving a separate expression for each of the two errors, leads directly to a form easily adaptable to further manipulation.

In totaling the squared class marks for all classes, the frequencies may be successively denoted by $f_a, f_b, f_c \cdots f_z$. These symbols may be substituted for f_{-1} and f_{+1} in the appropriate classes. In the first class it will be necessary to take f_{-1} as $-f_a$, and in the last class f_{+1} as $-f_z$, in order that the series of class parabolas may meet the condition that they should begin and end at $Y = 0$. The total for all classes may then be reduced to the form,

$$\sum fm_c^2 = \sum fm^2 - \left[N - \frac{4}{5}(f_a + f_z)\frac{k^2 + 1}{k^2}\right] \frac{i^2}{12} \cdot \frac{k^2 - 1}{k^2} \cdot \quad (6)$$

If the scale is considered to be continuous, and if subclasses are not appropriate, this formula obviously becomes

$$\sum fm_c^2 = \sum fm^2 - \left[N - \frac{4}{5}(f_a + f_s)\right]\frac{i^2}{12}\cdot \tag{7}$$

And if high contact is assumed, f_a and f_s disappear and the formula is in agreement with Sheppard's correction.

TABLE I

CLASSIFICATION OF CERTAIN SHOPS ACCORDING TO NUMBER OF MECHANICS EMPLOYED ($i=3$), WITH DIRECT COMPUTATION OF SUBFREQUENCIES IN UNIT CLASSES ($k=3$), AS DETERMINED BY PARABOLIC INTERPOLATION*

Class	Class mark m	Frequency f	Parameters of parabola	$k\times$sub-frequency kf_s	Subclass mark m_s	Summed squares $kf_sm_s^2$
			$a = 22.\underline{6}$	$9.\underline{3}$	1	$9.\underline{3}$
1– 3	2	20	$b = 9.\underline{3}$	$22.\underline{6}$	2	$90.\underline{6}$
			$c = -4.0$	28.0	3	252.0
			$a = 39.\underline{1}$	$33.\underline{7}$	4	$540.\underline{4}$
4– 6	5	36	$b = 0.\underline{6}$	$39.\underline{1}$	5	$977.\underline{7}$
			$c = -4.\underline{6}$	$35.\underline{1}$	6	1264.0
			$a = 24.0$	28.0	7	1372.0
7– 9	8	24	$b = -4.0$	24.0	8	1536.0
			$c = 0.0$	20.0	9	1620.0
			$a = 11.\underline{3}$	$15.\underline{3}$	10	$1533.\underline{3}$
10–12	11	12	$b = -3.\underline{0}$	$11.\underline{3}$	11	$1371.\underline{3}$
			$c = 1.0$	$9.\underline{3}$	12	1344.0
			$a = 5.\underline{7}$	$7.\underline{7}$	13	$1314.\underline{4}$
13–15	14	6	$b = -1.\underline{6}$	$5.\underline{7}$	14	$1132.\underline{4}$
			$c = 0.\underline{3}$	$4.\underline{4}$	15	1000.0
			$a = 2.0$	$3.\underline{3}$	16	$853.\underline{3}$
16–18	17	2	$b = -1.\underline{3}$	2.0	17	578.0
			$c = 0.0$	$0.\underline{6}$	18	216.0
				3)300.0		$3)17005.\underline{1}$
		$N=100$		$N=100.0$		$\Sigma f_sm_s^2 = 5668.\underline{370}$

Graphic height of subfrequencies ($x = m_s - m$)

$$Y_s = kf_s = f + \frac{2f - f_{-1} - f_{+1}}{8}\cdot\frac{k^2-1}{k^2} + \frac{f_{+1} - f_{-1}}{2i}x + \frac{3(f_{+1} + f_{-1} - 2f)}{2i^2}x^2$$

$$\Sigma fm = \Sigma f_sm_s = 662;\quad M = 662/100 = 6.62$$

$$\Sigma fm_c^2 = \Sigma fm^2 - \left[N - \frac{4}{5}(f_a + f_s)\frac{k^2+1}{k^2}\right]\frac{i^2}{12}\cdot\frac{k^2-1}{k^2}$$

$$= 5722 - \left[100 - \frac{4}{5}(20+2)\frac{10}{9}\right]\frac{9}{12}\cdot\frac{8}{9} = 5722 - 53.\underline{629} = 5668.\underline{370}$$

$$(N-1)\sigma_c^2 = \Sigma fm_c^2 - (\Sigma fm)^2/N = 5668.\underline{370} - (662)^2/100 = 1285.93\underline{0}37$$

$$\sigma_c^2 = 12.9892;\quad \sigma_c = 3.60$$

* Underscored figures are repetends.

Since it may be shown that $\sum fm$ requires no correction for slope, the estimated population variance is simply

$$\sigma_c^2 = \frac{1}{N-1}\left[\sum fm_c^2 - \left(\sum fm\right)^2/N\right]. \tag{8}$$

In order to illustrate the theoretical steps implied in parabolic correction for grouping, the accompanying table is presented, and the data and computed subclasses are charted. The data of the table may be assumed to represent a sample of 100 small shops classified according to the number of mechanics employed. The original data are grouped in classes of three; i.e., shops having 1 to 3 mechanics, 4 to 6 mechanics, etc. An estimate of variance is required, based on three subclasses in each class. It will be noted that the result obtained by the tedious estimate of probable subfrequencies and the totaling of squares of subclass marks, duplicates that obtained by the derived correction formula.

Very irregular distributions, such as bimodal, if similarly treated, would not always present a plausible smoothing in detail. In particular, when an extreme frequency (f_a or f_s) is relatively very small, parabolic interpolation may not be literally plausible, since a portion of Y in the class may be negative. But the use of the parabolic residual, f_a+f_s, in the correction formula, is plausible, since relatively small extreme frequencies, as in the case of the binomial $(0.5+0.5)^n$ with n large, imply an approach to the high contact which Sheppard's correction requires.

PRACTICAL APPLICATIONS OF THEORETICAL SAMPLING METHODS*

By John N. Webb, M. Starr Northrop,
and Stanley L. Payne
Bureau of the Census

In a recent popular novel, *The Just and the Unjust*, two attorneys are discussing the difference between the theory and the practice of law. One of the attorneys draws the distinction in terms that assume an extremely high level of professional competence: "Theory," he says, "is where you want to go; practice is how you are going to get there." The other attorney, denying that a high level of professional competence can be assumed, states the distinction in these terms: "Theory is what you tell people you are going to do; and practice is what they catch you really doing."

In this paper on some practical applications of sampling methods, it would be pleasant to take the first distinction between theory and practice as applying to statisticians as well as to attorneys. Were this as possible as it is desirable, we could take leave of sampling theory at this point with the remark that it provides us with the rules and the blueprints of our operations.

But some of you must have wondered from time to time whether statistical surveys in the service of marketing, or government, or opinion measurement, to mention only a few, can actually live up to the specifications laid down by sampling theory and still get the job done on time and within the limitations of field resources. When, for instance, you see a statement in a commercial report, a newspaper, or a monograph that the findings presented therein to one or two decimal places have an accuracy such that 95 or 99 times out of 100, the findings of the sample survey would not differ from the actual facts by more than some extremely small value, have you not sometimes marvelled at the high precision we statisticians have achieved in recent years?

Sampling theory does provide us with simple rules by which we can compute the reliability of our findings, provided certain procedures were actually carried out; and there is no lack of zeal in following theory on this point. The conditions under which the computation of reliability is valid are by no means easy to realize, particularly in sample surveys of social and economic phenomena. Moreover, practical adjustments in terms of field needs and dead-lines often complicate the problem of measuring the precision of the results, so that it is possible to make only broad statements about the reliability of the results.

---

* This paper was prepared for presentation at the 104th Annual Meeting of the American Statistical Association which was to have been held at Cleveland, Ohio, December 29 to 31, 1942.

Let us take, for example, the apparently simple requirement that every observation in the universe to be sampled must have an equal chance of being selected. Now, if our observations are numbered capsules in a selective service drawing, or cards in a file drawer, or colored marbles in an urn, this condition should be relatively easy to fulfill, although even then there may be unexpected difficulties. But if our universe consists of vacant dwelling units, or employed or unemployed workers, or passengers boarding intercity trains and busses, how are we to determine who or what makes up our total universe so that we can insure every item an equal chance of inclusion in our sample? To list all dwelling units in a community as a preliminary step to sampling is obviously ridiculous, especially when the listing would cost more than is allowed for the entire survey. Equally, to make a card index of all workers so that we may sample proportionately the employed and the unemployed is to take a census in order to take a sample.

Thus, the rules of sampling, however easy or difficult to devise when the universe is ordered and static, become extremely difficult of realization when the universe is disordered and dynamic. In many cases the theoretical conditions of sampling cannot all be met within the limits of time and cost, and resort must be had to procedures that we believe approximate the conditions desired. How far the sampler may safely depart from strict conformance to established rules is a matter for his competence and conscience to decide, because there is seldom any absolute check on the findings.

The sampler of social and economic data must improvise procedures from his knowledge of methods used in other fields and frequently must pioneer in developing new means of insuring sufficient reliability of results to meet administrative needs. Under the urgencies and inconveniences of the war program, he faces more and more difficult operating problems but by the use of ingenious sampling devices he continues to produce the necessary information. More than ever he must take practical considerations into account in devising methods without overlooking the necessary theoretical considerations. The concrete examples that follow show how theory and practice are brought together in two types of sample field surveys now in operation.

VACANCY SURVEYS

The urgent need for current information on residential housing for in-migrant workers in war industries provides the first illustration. Consider a small city near some war activity—a city in which the War

Production Board, the National Housing Agency and the War Manpower Commission are concerned about the amount of housing available for an expected influx of war workers and their families. Let us assume that activity is well under way and that already owners of residential property have found it their profitable and patriotic duty to convert large one-family structures into rebuilt apartments. Other landlords are breaking up their small apartments into separate rooms for rent. Thus, at the same time and for the same reasons, new small family dwelling units are being created from large residences and other small family dwelling units are being eliminated. It is difficult to determine in advance whether the family dwelling unit inventory is expanding or contracting under these circumstances.

Current information on dwelling unit vacancies is urgently needed in order that the city may be allowed its necessary number of public housing projects and that its private builders may receive their share of materials for new construction. The sampler is asked to design a method of obtaining the vacancy rate with a margin of error not too great for the purposes of the agencies administering the construction program. As usual, the findings are needed "day before yesterday" but they will still be of some value, let us say, four weeks from now. The agency requesting the data feels justified in spending less than one-fifth of the cost of one dwelling unit on this one survey. Furthermore, it reasonably expects the findings in this city to be comparable with those in more than 100 other cities where similar surveys are to be made, in order that the supply of housing materials may be distributed judiciously. Such are the conditions that face the sampler when he ventures to design the sampling method. Such conditions are not set forth in any elementary textbook nor in this particular combination even in advanced books on sampling theory. Since the sampling method he employs has never appeared in print it is up to the sampler to develop his own theory.

Let us suppose that the city in question is known to have some 12,000 dwelling units, that the expected vacancy rate is in the neighborhood of 2 per cent of all the units, and that the sampler is allowed a standard error of approximately 10 per cent of the reported vacancy rate. Reference to elementary probability theory will tell him that he could take a random sample of 3,500 units or about 3 in every 10 dwelling units in this city, but the theory does not indicate how randomness of selection is to be assured. An obvious device would be to select 3 in every 10 addresses from an up-to-date city directory. But even an excellent city directory cannot be kept up-to-date under wartime conditions; and if such a directory were available this device might still be impractical from the cost angle. Were it not for the time, the cost, and the problem of dwindling manpower, the sampler could make his own complete and current inventory from which the sample might readily be selected. Faced with the realistic problem of providing the information in less than a month at a cost of somewhat under $1,000 and having no current city directory in which he places reliance, our sampler instructs his local supervisor as follows:

He is to obtain the most recent map of the city and bring it up to date by conferring with the city engineer on new streets which have been cut through and boundaries which have been changed. He is then to subdivide the map into areas as homogeneous as possible, using as a first approximation Census tracts, wards, school districts or other recognized local divisions. Then he calls upon one of the city fathers to assist him in further subdividing these areas into white and Negro neighborhoods, the rooming house section and other important types of area. If the supervisor had the time, he might make a scientific stratification using the Real Property Inventory, for example, but the sampler explains that the old resident's idea will serve the purpose almost as well. The supervisor is then to number seriatim all the blocks within each sub-area, carrying over the numbering from one sub-area to the next. Finally, he is to select the sample blocks by simply taking every other one of the numbered blocks. This sampling ratio will, of course, give a larger number of dwelling units than would have been covered in the hypothetical 3 in 10 random sample prescribed by the textbook; but more important, it gives a smaller number than would have been included in a complete inventory of the city.

Next, the local supervisor proceeds to have his enumerators make an inventory of dwelling units in these sample blocks, ringing each doorbell, and investigating all alleys and passageways in their search for all units whether in converted attics, in basements, or over garages. As the enumerator makes the inventory of a given block, he lists the pertinent information for each unit which he discovers to be vacant (an average of only one vacant unit in every 50 according to the original assumption), listing the number of rooms, the rental asked, the condition of the structure and the facilities included in the dwelling unit. Thus, when he has completed his inventory of a sample block, he has also completed the count and description of the vacant units within that block, at little more than the cost of a simple listing of addresses.

Experience has taught the sampler that the block sample, which is an areal sample drawn from a horizontal plane, might, by chance inclusion of too many large apartments rising above the horizontal plane, seriously increase the apparent vacancy rate of the sample over the true vacancy rate. This is especially true of rates for rental vacancies since the large apartments frequently have high turnover of tenants and because all the apartment units are rental units. To overcome this difficulty, the local supervisor is asked to exclude the large apartments from the block sample and to investigate the apartments in *all* large apartment buildings whether they fall in the block sample or not. This is relatively easy because some one person in the apartment house, the manager or the janitor, can usually give all the necessary information on the vacant units. Results for the block sample are weighted up by the sampling ratio and results for the large apartment structures are added in to give the combined vacancy rate for the city. In this way the administrative conditions are met within the bounds of sound theory. Each block has an equal chance of representation in the sample; densely populated and sparsely populated areas are represented in the same proportions.

The important decision in using the block sample method is the estimate of the efficiency of the block as a sample unit in comparison with that of a random sample of addresses. This estimate is difficult to make for the first survey in a given city, but the approximate number of blocks, the number of dwelling units, and the April 1940 vacancy rate are helpful in typing the city for this purpose. For use in later surveys of this city investigation of the results of the first survey by blocks gives an excellent estimate of the efficiency of the block sample. A great boon to the sampler in this connection are the tabulations of block statistics from the 1940 Census now available for all cities over 50,000 population. Not only is it now possible to use a locational stratification but also to employ other means of stratification, such as race of occupants, age of structure, and average monthly rental.

LABOR MARKET SURVEYS

The sampler's problems are not limited to designing plans for housing surveys alone. In this period when the Nation needs every available man and woman to serve on either the front line or on the production line, timely statistics on manpower are needed on a current basis for use in planning the war effort and in relieving labor shortages. How many civilians are now working? At what occupations and industries? Can any of these be shifted to more important jobs? What is the size of the unemployed labor reserve? Who are the unemployed and what can they do? Is the present labor force large enough to do the job ahead, and if not, what is the labor reserve of nonworkers who can and will take jobs if the need arises? What are the work experience and skills of persons in this labor reserve; will they need training and what kind?

If married women come into the labor market, will day nurseries have to be established?

Some or all of these questions require immediate answers for the guidance of persons dealing with the manpower problem. Results may be required for a given city, a county, a local labor market area, a State, a geographic region or the Nation as a whole. A complete census of the United States could provide answers to all of these questions for all types and combinations of geographic areas, but as in the case of the housing surveys the results are wanted day before yesterday. Information about employed workers might be obtained by a canvass or questionnaire directed to all employers or a sample of employers.

At this point, however, the sampler begins to ponder the problem of establishing his universe of all corner grocers who hire help, of all housewives who hire servants, of all home owners who hire the neighbor boy to mow the lawn, and so on. Even if he solves these problems, and he probably could, he immediately realizes that he would not be able to answer all of the questions raised above. The employer cannot tell how many unemployed there are in his community, nor can he furnish any information on how many housewives or high school students would take full-time jobs in industry. It soon becomes apparent that, to be complete, a labor force study must include information on all types of persons—employed, unemployed, and nonworkers. As mentioned above employers' records would provide at least a partial listing of employed workers, but there is no such list of the unemployed in any city, county or State. Work relief agencies and employment offices could provide partial lists of the unemployed, but these would be incomplete. Outside of school records, there are practically no listings of nonworkers.

At this point the sampler has ruled out the complete census as being too costly and time-consuming, but he has also discovered that there are no tailor-made lists of his employed, unemployed, and nonworker universes from which he can sample. As was pointed out previously, to prepare complete lists or card indexes would be too costly and time-consuming. Regardless of whether the study is to cover a city, county, State, or the whole country, the required data must be obtained from a direct enumeration of a cross-section of the population within the primary sampling unit. The "Monthly Report on the Labor Force" survey, conducted each month by the Bureau of the Census,[1] covers a national sample in which the primary sampling units are counties and the secondary sampling units are households (dwelling units). The procedures followed in selecting the subsample of households within these counties serve to illustrate some of the practical approaches to problems of sampling for labor market data.

---

[1] See *Sampling Procedures and Method of Operation of the Monthly Report on the Labor Force*, U. S. Department of Commerce, Bureau of the Census, November 1942.

The sampler has worked out a carefully designed stratified sample of counties. He has been informed of the phenomenon that is to be measured; and taking into account funds available, field operating problems, and estimates of the sampling error, he has determined the total number of observations to be made. The total number of observations are broken down into county quotas. In turn, the quota for each county is assigned to open country areas, cities, towns, etc., in proportion to the total population of the county residing in those places.[3] So far, so good —all of the preliminary work has been painstaking and very important, but it has all been done by an "expert IN Washington." Now comes the job of deciding how a local supervisor and finally an enumerator out in county *X* will locate and interview a sample of, say, 300 households that will represent all types of situations according to the detailed quotas established for that county. Households must be sampled from among those prosperous farms at the northern part of the county; the poorer farmers in the southern end must also be represented; quite a large quota is in "the city"; those five or six small incorporated towns must be covered; and finally the unincorporated towns—"the wide places in the road"—cannot be overlooked.

In cities and towns for which maps can be obtained the local supervisor proceeds as follows:

1. An accurate map of the city or town is obtained, checked, and numbered in the same manner as outlined above for the housing survey.
2. A block quota is obtained by dividing by three the total number of households to be visited. This means that on the average three calls will be made in each block, a consideration that is important in cutting down travel time and cost, and also in saving gasoline and tires.
3. The block quota is then divided by the total number of blocks in the city. If this comes out to 20, then every 20th block in the city is designated as a sample block. This guarantees geographic coverage of all sections of the city.
4. Block descriptions, in terms of street boundaries, are then transcribed from the map to block lists.

---

[3] Another plan for assigning a quota within the primary sampling unit (e.g., county) of a stratified sample is to make allocations in such a way that the observations will represent the population of the entire stratum from which the primary sampling unit was selected, rather than the population of the primary unit alone. To cite a simple example, assume county *X* is selected to represent a given stratum, and that 20 per cent of the population of this county lives on farms. Assume further that for the total population of all counties in this stratum, 25 per cent of the population lives on farms. Then for this county 25 per cent, rather than 20 per cent, of the county quota would be assigned to farm households.

5. Enumerators are sent to each sample block, and on a separate line of the block list form they record the address of every dwelling unit in the block, whether occupied or vacant.
6. When all block lists are complete and checked, a count is made of the total number of dwelling units in the total of all sample blocks.
7. This count is then divided by the number of dwelling units to be sampled in this city to provide a dwelling unit ratio.
8. The dwelling unit ratio is then applied to the block listings of addresses. If the dwelling unit ratio is 10, then the 10th, 20th, 30th, 40th, etc., dwelling units listed become the households to be included in the sample. By this device, blocks that contain only a small number of dwelling units may have possibly one or no addresses included in the sample. On the other hand, congested blocks in tenement or apartment house districts may have 10 or 15 addresses designated for the sample.

In small towns and unincorporated places for which no maps can be obtained the local supervisor proceeds on the basis of one of the following alternatives:

1. He draws a rough draft map showing street names and other land marks, and proceeds as outlined above; or
2. He omits steps 1 through 4 as outlined above, has his enumerators block list *all* blocks in the town; and proceeds with steps 5 through 8.

It should be noted that the block lists as described above should be carefully filed after the sample is drawn. These provide an immediate source of additional addresses from which other samples can be drawn without further expense. In general, however, block lists that are a year or more old should be brought up to date by a canvass of the blocks so that newly constructed units can be added and eventually sampled.

For open country areas, including very small unincorporated places or cross-road clusters of dwelling units, the sampler proceeds as follows:

1. The total open country quota for the county is divided up (in units of 5) between townships or similar minor civil divisions in proportion to the open country population living in each.
2. A county map showing roads and highways is obtained. If the pattern of roads does not already break up the map into mile square sections, it is ruled in pencil vertically and horizontally at 1 mile intervals to accomplish this same result.
3. Within each township these sections are numbered consecutively.
4. For every 5 units in the quota for a given township, 3 sections are

designated for inclusion in the sample by taking every *n*th section. For each section so designated, a starting point is marked arbitrarily with an *X*.

5. At this stage, the map is sent to the field for use by the local supervisor.
6. To determine the open country sample for a certain township in the county, the supervisor proceeds to the lowest numbered section on the map, locates the starting point, and traveling around the section in a counter clockwise direction, writes a descriptive address for each of the first five dwelling units (either side of the road) that he comes to. If a section does not provide 5 households, he proceeds to the section with the next higher number, etc., until the quota for the township is reached.

The procedures illustrated immediately above for labor market surveys are based on the assumption that fixed schedule quotas can be assigned to the primary sampling units. This is usually possible where current population figures are known. At the present time, however, population migration as a result of war and other conditions has become so pronounced in certain areas that the sampler will be hesitant to establish fixed quotas on the basis of population figures that do not reflect recent increases or decreases. This problem may be met by establishing fixed sampling ratios rather than quotas. For example, a town having 1,000 households in 1940 might warrant a sample of 50 households or 1 in 20. Thus the quota might be established by this rule: take every other block and every 10th dwelling unit in each sample block. If there has been no population change the sample should be close to expectations, i.e., 50 households, but if the population has doubled since 1940 this ratio quota should provide a sample of 100 households. Ratio quotas can be adapted to the techniques described above with very few changes.

The actual application of sampling designs as described above for housing and labor market surveys illustrates only a few of the problems that must be faced in connection with the planning and operation of sample surveys. The sampler must foresee and take account of practical field problems in arriving at a sampling plan. Regardless of how nearly perfect a sampling design is from a theoretical standpoint, it must be discarded if the field staff cannot take over actual operation on a practical basis, and within necessary time and cost limits. Both theoretical and practical considerations must be given full weight; there will often be compromises between the two, but neither should be disregarded in developing a sound theoretical sampling plan that can be put into operation on an efficient basis by nontechnical field supervisors and enumerators.

SOME SAMPLING USES OF DATA FROM THE CENSUS OF AGRICULTURE*

By Irvin Holmes†
Bureau of the Census

The 1930 census of Agriculture provided, among other things, a classification of farms into 13 major types and 5 subtypes. This classification by type of farm was based primarily upon the value of farm products sold, traded, or used by the operator's family. It was used for making type-of-farm tabulations by counties and states of many agricultural items, such as acreage of land in farms, value of farm land and buildings, receipts from boarders, lodgers, etc., numbers of specified kinds of livestock and livestock products, numbers of farms by size, numbers of farms by value of products, numbers of farms by tenure of operator, and expenditures for specified items, as well as the value of farm products. As a result of the experience gained in this work, changes were made in the county and state tabulations of value of farm products data for the 1940 Agriculture Census as follows: The number of questions on value of farm products was increased from 6 to 10 to provide more detail regarding sources of income; classification of farms by major (single largest) source of income was substituted for classification by type in order to provide a more objective basis for coding and, thereby, reduce editing costs; and the value-of-farm-products data were tabulated for farms classified by total value of farm products in order to show the principal sources of income of farms at different gross farm income levels.

County and state tabulations were not made for farms classified according to major source of income, or according to total value of farm products, for the additional agricultural items mentioned earlier in connection with the 1930 Agriculture Census tabulations. Instead it was decided to use sample data as a basis for presenting statistics on the characteristics of farms classified according to total value of farm products and according to type (major source of income). Four considerations contributed mainly to this decision: The urgent need for quickly available figures in connection with war programs, particularly statistics on marginal (low income) farms and on farm labor; the necessity for holding costs to a minimum; the greater variety of tabulations that

---

* This paper was prepared for presentation at the 104th Annual Meeting of the American Statistical Association which was to have been held at Cleveland, Ohio, December 29 to 31, 1942.

† The author acknowledges the valued assistance rendered by Morris H. Hansen, William N. Hurwitz, and Hilton E. Robison of the Bureau of the Census.

could be made on a sample basis than on a complete basis with the funds available; and the possibility of securing experience and information which would be valuable in preparing plans for future censuses, either on a complete or on a sample basis. The purpose of this paper is to review the work that has been done to date on these tabulations of sample data from the 1940 Census of Agriculture.

SAMPLE DESIGN AND TYPE OF SAMPLING UNIT

The primary objective of these tabulations of sample data was to secure national statistics on the characteristics of farms classified according to total value of farm products, according to major source of income, and according to color and tenure of operator. The secondary objective was to provide measures of the variation in these statistics by large geographic regions. All of the data for one farm for a group of related items, such as value of farm products, are on a single punch card. Consequently, it was feasible to use the smallest available sampling unit, that is, the individual farm. Most previous studies of sampling techniques for proposed sample censuses of agriculture have centered about the problem of size of sampling unit; i.e., the problem of balancing enumeration difficulties and costs for small sampling units, such as individual farms, against the decrease in precision resulting from the use of larger sampling units, such as clusters of farms.

The number of farms, according to the 1940 Agriculture Census, varies from 135,190 for the New England Division to 1,090,574 for the West North Central Division. It was recognized that, under these conditions and with the objectives as stated, it would be highly desirable to vary the sampling ratio of number of farms for each geographic area for which statistics were desired. Unfortunately, administrative considerations made this procedure impractical. It was also recognized that for tabulations of farms classified by major source of income, the sampling ratio should be varied by major-source (type) groups. Likewise for any tabulations of farms classified by total value of farm products, as well as any tabulations by other sorts which involved value-of-farm-products data, it would be desirable to vary the sampling ratio of number of farms by value of farm products groups. The same administrative limitations prevailed here; consequently, the following procedure was substituted:

The chief interest was in tabulations involving farms classified by total value of farm products. The farms with $10,000 and over total value of farm products represented only 1 per cent of all farms, but accounted for approximately 17 per cent of the gross farm income for the United States in 1939. Accordingly, the cards for all farms with $10,000 and over total value of farm products were sorted out and tabulated completely.[1]

The 1940 Farm and Ranch Schedules had been numbered by machine consecutively within each county prior to the time that the data were placed on punch cards. Before numbering, the schedules were arranged by minor civil divisions within each county. For the most part, the schedules were in order by enumeration districts within each minor civil division; and to a considerable extent, the schedules for each enumeration district remained in the original order of enumeration. The punch cards for the sample farms were selected by machine by sorting for each county all cards with serial numbers ending in 15 and 65. The same terminal digit was selected to reduce the number of cards handled in the sorting. The use of these serial numbers excluded 22 counties from the sample and resulted in an oversampling of the counties with small numbers of farms, particularly those with less than approximately 100 farms each. If serial numbers ending in 25 and 75 had been used, the oversampling of the small counties would have been reduced, but the number of counties excluded from the sample would have been increased to 28. For a national and a regional sample, neither of the above defects appears important. At the time of the 1940 Census of Agriculture, there were only 64 counties with less than 100 farms each, or approximately 2 per cent of the total number of counties in the United States. Furthermore, of the 22 counties excluded from the sample, 10 were independent cities in Virginia.

The sample drawn has, therefore, a primary stratification into 2 classes by total value of farm products per farm in 1939. For farms with $10,000 and over total value of products, a 100 per cent sample of farms was taken. For farms with less than $10,000 total value of farm products, a 2 per cent sample of farms was selected. The secondary stratification for the latter group was geographic, that is by counties and, to a considerable extent, by minor civil divisions within each county. Furthermore, to the extent that the original order of enumeration was preserved in numbering the schedules, the 2 per cent sample of the low income farms approximates a selection along the enumerator's route of every 50th farm in that category.

SCOPE OF TABULATIONS

At the close of 1942, the following tabulations on a sample basis were completed or in progress:

---

[1] Actually, sample cards were also selected for farms with $10,000 and over total value of farm products. For certain tabulations, these sample cards have been used instead of the complete tabulations because of other considerations.

(1) Tabulations of numbers of farm laborers, expenditures for farm labor and other specified items, acreage of land in farms, value of farm land and buildings, work off farm, age of operator, year of occupancy, farm machinery and facilities, cooperative buying and selling, and numbers of specified kinds of livestock for farms classified by total value of farm products, 1939.

(2) Tabulations of acreage of land in farms, value of land and buildings, farm machinery and facilities, work off farm, and specified farm expenditures for farms classified by major source of income, 1939.

(3) Tabulations of value of farm products sold, traded, or used by farm households, 1939, by sources of income; and total value of farm products by value groups and by major-source groups for farms classified by color and tenure of farm operator.

The tabulations described under (1) were being made in cooperation with the Bureau of Agricultural Economics and the Farm Security Administration of the Department of Agriculture. State tabulations of these data have been furnished to these agencies for administrative use in connection with problems relating to farm labor and marginal farms. Preliminary reports, or press releases, are also being prepared showing figures for those items by geographic regions for which the expected error is small enough to justify publication. Present plans are to publish the results of all three of these series of tabulations in a special monograph.

SOME COMPARISONS ON COSTS OF SAMPLE TABULATIONS AND COMPLETE TABULATIONS

The cost of pulling and coding the 2 per cent sample and making one tabulation run *by states* has been estimated, for this work, at roughly one-sixth to one-eighth of the cost of making a single tabulation run *by counties* for all farms. In both cases, the cost of reproducing the codes from one card to another has been included; and the cost for the sample tabulation includes the complete tabulation for the farms with $10,000 and over total value of farm products. The cost of coding and tabulating these large income farms exceeds the cost of pulling, coding, and tabulating the 2 per cent sample cards for farms with less than $10,000 total value of farm products. The cost of making an additional run of the sample cards *on a state basis*, after the sample cards have been drawn and coded, is estimated in the neighborhood of 1 per cent of the cost of making a single tabulation run *by counties* for all farms. All of these comparisons relate only to costs for machine tabulation and do not include costs of preparing the data for publication. The technical and clerical costs of processing these sample data have proved to be materially higher, on a per unit basis, than the corresponding costs for mass processing of complete tabulations. Additional experience will probably result in lowered clerical costs on this phase of the work; but, if the Bureau of the Census is to render extensive service in this type of work, an adequate technical staff should be provided.

METHODS OF EXPANDING THE SAMPLE

For tabulations (1) and (2) previously described, two factors were available for expanding or inflating the sample. Published figures were available for the total number of farms and total value of farm products for each value-of-farm-products group[2] and also for each major-source-of-income group. For the tabulations outlined under (3), only one expansion factor was available since the only known values were the total numbers of farms in each color-tenure group. For the tabulations under (1), it was decided to use the total number of farms in each value group for expanding the sample data; that is, the sample figures for each value-of-farm-products group were multiplied by the reciprocal of the sampling ratio for that value group.

For nearly all of the items sampled, the totals for all value-of-farm-products groups were also known.[3] On such items, the expanded sample figures for each value-of-farm-products group were scaled to equal these previously published totals by simple proportionate adjustments. In some of the tabulations, cross classification tables were involved, such as numbers of farms cross classified by total value of farm products and by size of farm. In such cases, where both the column and line totals were known, the expanded sample figures for all cells were adjusted to equal these known totals by the process of iterative proportions. That is, a succession of horizontal and vertical proportionate adjustments were made until the frequency table became rigid. These adjustments were made for the convenience of any user who might wish to relate the figures to previously published Census data.

It was recognized that for some of the tabulations under (1), particularly such items as all land in farms, value of land and buildings, and expenditures for specified items, some increase in precision might have been obtained had the sample data been expanded by using an inflation factor based on the total value of farm products for each value-of-products group. The decision to use total number of farms in expanding the sample was determined partly by administrative reasons and partly by the fact that, since the tabulation sort was by value of farm products, much of the over-all relationship between value of farm products and such items as value of land and buildings and farm expenditures had already been taken into account.

---

[2] For two value-of-farm-products groups the unclassified farms and farms with no farm products sold, traded, or used by farm households, only the total number of farms was available.

[3] The same is true for item totals for all major-source groups and for all color-tenure groups. The discussion in the remainder of this paper relates only to the tabulations outlined previously under (1), since only the figures for these tabulations were available at the time this paper was written.

A third expansion factor, which could have been used, was the reciprocal of the sampling ratio based on number of farms for the total sample, that is, for all value groups combined. Obviously this would have given less precision than that obtained by the use of sampling ratios for the individual value groups. This would be particularly true for the terminal value groups where the population numbers of farms are relatively small and the sampling ratios quite variable.

SAMPLING ERRORS FOR SOME SPECIFIED ITEMS

The first tabulations completed under (1) involved principally proportions of the total number of farms in each value group reporting specified characteristics. The exceptions were such items as acreage of all land in farms, acreage of cropland harvested, value of land and buildings, age of operator, and number of days of work off farm.

It is well known that when the number of farms in the sample is large, the relative sampling error (coefficient of variation) will be small. Thus, in the central value groups, for which the sample number of farms is large, it is to be expected that the relative sampling errors will be considerably smaller than those for the terminal value groups in which few farms are represented. Hence, for these terminal value groups, the figures presented for some of the geographic regions may be subject to considerable error, where based on sample data rather than complete tabulations. For example, there is some evidence that, in order to obtain a 10 per cent relative error on cropland harvested for the "$20,000 and over" value-of-products group for the Pacific Division, it would be necessary to increase the sample size from 2 per cent to about 5 per cent; while for all land in farms, the same relative accuracy could be obtained only with a sample of approximately 15 per cent.

The relative errors to be expected differ as between items. For example, in estimating age of operator, the size of sample necessary for a given relative error is considerably less than that required for all land in farms, cropland harvested, value of land and buildings, and work off farm. Some evidence is available from California and Illinois:

For the "Unclassified farms" and for the "Farms with no products" in Illinois, it was found that, to secure the same relative error on all land in farms and value of land and buildings as for age, it would be necessary to obtain samples 6 to 8 times as large. The same was true of value of land and buildings and work off farm for California farms in the "Farms with no products" classification. On the other hand, for the "Farms with no products" group, it would be necessary to increase the size of sample on cropland harvested nearly 50 times in Illinois and about 90 times in California to obtain the same relative accuracy as for average age of operator. The farms in this no-income classification include not only small part-time farms near villages and cities, but also some farms with large acreages but no income because of crop failure or other unusual conditions, which accounts for the large variability.

For the "$10,000 to $19,999" value-of-products group in Illinois and the "$10,000 to $19,999" and "$20,000 and over" value groups in California the sample size for all land in farms would need to be from 40 to 80 times as large as that for age of operator to secure the same relative accuracy. For cropland harvested and value of land and buildings for these value groups the indicated increase in sample size to secure the same precision as for average age ranges from fivefold to twentyfold. The farms with large total acreages in the higher value groups include many with a high proportion of the total area made up of pasture land, woodland, and rough land, rather than cropland. Such farms have a relatively low value per acre. On the other hand the small size farms in these upper value groups are hatcheries, greenhouses, nurseries, and other highly specialized operations for which the average value of land and buildings per acre is comparatively high. The problem of the extreme variability in size of farm, particularly in the western states, was partly met by making complete tabulations of the acreage of all land in farms, cropland harvested, and value of land and buildings for the very large size farms, generally those of 30,000 acres and over.

For work off farm in the highest value-of-products groups the sample data for California indicate that it would require a sample approximately 250 times as large as that for average age of operator to secure the same relative precision. In these value groups the frequency for number of days operator worked off farm is distinctly bimodal. Most operators of large-income farms report very little or no work off farm, but a few such operators report nearly full-time nonfarm employment, generally at professional or business occupations. To secure statistics on work off farm for these value groups with a relative error of 10 per cent would require approximately a 33 per cent sample, which figure includes a correction for sampling from a finite population.

The table shows the coefficients of variation by value-of-farm-products groups for three of the farm characteristic items for the East South

TABLE I

PER CENT COEFFICIENTS OF VARIATION FOR PROPORTION OF TOTAL FARMS IN EACH VALUE-OF-FARM-PRODUCTS GROUP REPORTING SPECIFIED CHARACTERISTICS, FOR THE EAST SOUTH CENTRAL AND PACIFIC DIVISIONS

Division and value group	Number of farms*	CV for proportion of total farms in each value group—		
		Operated by full owners	Reporting no work off farm	Reporting horses and/or mules and no tractors
		Per cent	*Per cent*	*Per cent*
EAST SOUTH CENTRAL DIVISION				
Unclassified farms	3,888	18.6	13.8	13.3
Farms with no products, etc.†	10,684	7.4	7.9	13.1
$1 to $99	62,979	3.0	2.8	4.1
$100 to $249	218,150	1.7	1.8	1.4
$250 to $399	222,481	1.8	1.2	1.0
$400 to $599	207,353	1.9	1.0	0.9
$600 to $749	93,290	2.7	1.4	1.8
$750 to $999	84,093	2.7	1.4	1.4
$1,000 to $1,499	64,299	2.8	1.7	1.6
$1,500 to $1,999	22,778	4.0	2.8	2.9
$2,000 to $2,499	10,720	5.4	4.2	4.6
$2,500 to $3,999	12,701	5.8	3.4	4.7
$4,000 to $5,999	5,175	8.1	5.5	10.3
$6,000 to $9,999	2,963	12.3	7.4	18.7
$10,000 to $19,999‡	1,310	21.0	10.1	46.4
$20,000 and over‡	485	105.6	23.2	70.2
PACIFIC DIVISION				
Unclassified farms	1,501	16.6	19.4	32.0
Farms with no products, etc.†	8,737	5.1	9.9	18.7
$1 to $99	19,575	2.5	6.3	11.4
$100 to $249	32,449	1.9	5.3	7.4
$250 to $399	27,578	2.2	5.4	6.9
$400 to $599	26,492	2.3	5.2	6.5
$600 to $749	14,900	3.2	5.8	7.4
$750 to $999	19,637	3.1	4.6	6.1
$1,000 to $1,499	28,491	2.5	3.1	4.9
$1,500 to $1,999	19,575	3.8	3.7	6.3
$2,000 to $2,499	14,259	4.5	4.0	7.5
$2,500 to $3,999	24,899	3.7	2.8	6.8
$4,000 to $5,999	14,726	6.0	3.1	10.1
$6,000 to $9,999	11,521	8.3	2.9	16.4
$10,000 to $19,999‡	7,654	10.2	2.8	17.8
$20,000 and over‡	4,179	16.0	6.0	25.9

* The sample size approximates 2 per cent of the numbers shown in this column.

† Farms with no farm products sold, traded, or used by farm households.

‡ For these items, sample data only were available for farms with $10,000 and over total value of products.

Central and Pacific Geographic Divisions. The East South Central Division, which is next to the largest geographic division from the standpoint of number of farms, includes Kentucky, Tennessee, Alabama, and Mississippi. The Pacific Division, which is next to the smallest geographic division in number of farms, includes California, Oregon, and Washington. In the East South Central Division the farms are concentrated in the smaller-income groups; in the Pacific Division the reverse is true. For these items, as well as for part of the other tabulations listed under (1), the sample cards for the farms of $10,000 and over were used instead of the complete tabulations. This procedure was necessary because the complete tabulations for the large-income farms could not be completed in time for use in these first studies.

This table illustrates the difficulty of sampling characteristics of farms in the terminal value groups where the numbers of farms are small. In the East South Central Division the sample indicates that full owners comprise only a very small proportion of the operators of farms with total value of farm products of $20,000 and over. To secure this statistic with a relative error of 10 per cent would require a sample of approximately 20 per cent of the farms in that value group. For the proportion of farms reporting no work off farm, a characteristic more frequently reported in this value group, it would require a sample of about 5 per cent to secure a relative error of 5 per cent, or a sample of about 25 per cent to obtain a relative error of 2 per cent. For the same value group it would be necessary to obtain a sample of about 30 per cent in order to secure a figure on proportion of farms reporting horses and/or mules and no tractors with a relative error of 5 per cent; if the desired precision were 2 per cent, a sample of 70 per cent of all farms in that value group would be needed. These comparisons are based on calculations which make allowance for the fact that the sampling is from a finite population. However, the formula used is for an unstratified sample; therefore, the coefficients of variation shown in the table may slightly over-state the error.

The sampling errors discussed here apply only to the figures for individual cells. However, the primary purpose of these tabulations is to determine patterns of relationships. The tabulations outlined under (1) were made on a state basis, although for most, if not all, items publication will be restricted to figures for geographic divisions. The consistency between states in the patterns of the various relationships justifies more confidence in the results, when used for an analysis of relationships, than might be inferred from the relative errors calculated for the individual cells.

MARKETING AND SAMPLING USES OF POPULATION AND HOUSING DATA*

By A. Ross Eckler and E. P. Staudt
Bureau of the Census

The 1940 census provides many types of population and housing statistics useful for the establishment of marketing programs, and in particular for the effective operation of sample studies required for market analysis. These data have much significance even in a period of sharply contracted output and distribution of civilian goods. It is now more important than ever that goods be marketed as efficiently as possible so that the limited supply can be sold at the lowest cost and at the least inconvenience to customers.

Statistics on the size and characteristics of the population and on the characteristics of the nation's homes are of basic importance to marketing men, since people are the buyers and consumers of the great bulk of goods produced for sale. Such data as the age and sex composition of the population and the tenure and rental value of homes are widely used as market indicators. It may be useful to mention briefly certain respects in which some of these uses have been modified during the war period, but it is not necessary to describe them here since they have been discussed in considerable detail elsewhere. Consequently, the greater portion of this paper will be devoted to a description of two types of use that have received relatively little attention hitherto but may become of much importance: (1) To describe the market with which a particular business organization is in contact, and (2) to provide area and other controls needed for special sampling studies of markets.

SIZE AND CHARACTERISTICS OF THE POPULATION

There are at least two ways in which the changes arising from war have affected the relative importance of data on the size of the population and of those on its characteristics. First, data on the size of the population have increased in relative importance because the interclass differences in buying patterns have been greatly reduced. This reduction is a result of the disappearance of many types of durable and luxury goods and the rationing and price control of certain other commodities. At a time when so large proportions of the population have sufficient purchasing power to buy more goods than they can get, it is less necessary than usual to formulate marketing policy in terms of variations in economic status. The sales of sugar, coffee, and gasoline are now being made on either a population or a need basis, and respond little, if at all, to changes in the purchasing power of various groups. Any further increase of rationing programs will expand the zone within which the volume of sales is almost directly proportional to the size of the population.

---

* This paper was prepared for presentation at the 104th Annual Meeting of the American Statistical Association which was to have been held at Cleveland, Ohio, December 29 to 31, 1942.

Second, data on the size of the population have particular significance when size is changing rapidly. The tremendous scale of internal migration revealed by recent estimates of the civilian population (based mainly on sugar registration figures) calls for a certain amount of revision in distribution programs, in the setting of sales quotas, and in the allocation of salesmen. These revisions should be made, however, with full regard for past population trends. If further revisions of market programs need to be made because of decreases in the amount of available goods or in number of personnel obtainable, every effort should be made to favor areas where past population trends suggest the likelihood of the continuation of expanding markets after the war.[1]

HOUSING DATA AS INDICATORS OF MARKETS FOR PARTICULAR KINDS OF GOODS

Data from the Housing Census can be used to some extent as direct measures of the potential markets for certain types of goods. Thus, information on specific types of equipment, such as mechanical refrigerators, radios, and central heating plants, indicates the number of dwelling units without a particular type of facility and thus the number of families in the untapped segment of the market. Information on the number of homes with electric lighting indicates the total number of families to which electrical equipment might be sold.

By the use of housing data on facilities in combination with data on population characteristics it is possible to determine the area, color, and tenure groups that form the largest potential markets for specific kinds of equipment. Most of these uses are of purely theoretical value at present since products of these types are not being produced and the supply still available for sale is negligible, but such marketing uses of housing data will become very important in the postwar period.

One exception to this generalization results from the rationing of fuel oil. Information on the number of homes having oil burners has been used as an indication of the potential market for various kinds of alternative heating devices, such as coal burners and fireplace grates, as well as for conversion equipment, and for insulating materials and other means of reducing heat loss. Housing Census data on home equipment may increase in importance if prolongation of the war economy requires a marked expansion in repair services.

---

[1] This subject is discussed in a paper entitled "Population Shifts and Income Changes," by Dr. Philip M. Hauser, Assistant Director, Bureau of the Census, presented on January 15, 1943, at the Wartime Marketing Conference of the American Management Association, in Chicago, Illinois.

HOW CENSUS DATA CAN BE USED TO SHOW CHARACTERISTICS OF PARTICULAR GROUPS OF THE POPULATION

Restrictions on the disclosure of data collected by enumerators make it impossible for the Bureau of the Census to furnish any consumer of statistics with information relating to a particular person or organization. The Bureau may, however, provide data in the form of totals pertaining to groups of people or organizations, provided a number of persons or firms are included. On this basis, it is possible to derive directly from census schedules compilations of the social and economic characteristics of groups of individuals that may be of much importance to a business organization—groups such as the subscribers to a magazine, the charge-account customers of a store, the patrons of a public utility, and similar groups of customers or subscribers for whom names and addresses are available for identification.

Statistics on such groups can be used in various ways in the marketing program of a particular organization. For example, data on the characteristics of the subscribers to a magazine may be very useful to the advertising department since many advertisers normally concentrate their expenditures upon publications that reach people in particular economic or social strata. Data on subscribers might be collected, of course, by direct field canvass or by mailed questionnaires, but both methods involve a considerable amount of planning and are subject to certain limitations. Information already collected on the decennial census schedules is much more comprehensive than any that can be readily obtained through field canvass, and has the merits of uniformity of procedure, completeness of coverage, and full cooperation on the part of the respondent. Moreover, data can be obtained in this manner from census schedules at a cost far below that necessary for the collection of similar information through field interviews.

The possibilities of this procedure may be illustrated by a brief description of several jobs which have been done on a contract basis for publishing firms. Two of these undertakings were carried out for the Curtis Publishing Company in connection with an analysis of a sample of subscribers to either the *Saturday Evening Post* or the *Ladies' Home Journal* and of a similar sample for the *Country Gentleman*. Under the procedure used, sample lists of subscribers to these magazines were sent to the Bureau of the Census for matching with names and addresses on the census records of 1940. Because of population movements since

1940 and other reasons, some of the names could, of course, not be located on the census schedules. For all those that were located, all available information was taken from the population schedule, the housing schedule, and (for persons living on farms) from the farm schedule. The tabulations of data thus assembled furnish a comprehensive picture of the economic and social characteristics of subscribers to the particular magazines, without disclosing any facts pertaining to a given individual.

A similar job has been done for the Meredith Publishing Company, the publisher of *Successful Farming* and *Better Homes and Gardens*. The conduct of this job was in general similar to that for the Curtis Publishing Company, although there were many differences in detail. The work entailed the selection of a sample list of subscribers to *Successful Farming*, the matching of this list against the farm schedules, and the transcription of selected agriculture, population, and housing data.

On the basis of such materials, publishing companies can determine the extent to which their subscribers are typical of the general population and in what respects their distribution with respect to such items as education, income, rent, and housing indicates that the subscribers are a specially selected group. Editorial as well as advertising policies may advantageously be adjusted to findings with regard to the composition of the subscriber-group.

The technique of using population and housing census files with their records of nearly three billion facts about people and the homes they live in seems to have considerable promise for recurrent use as a means of revealing the characteristics of particular groups. It seems likely that the reopening of old markets and the development of new ones after the war may well lead to a very substantial increase in the opportunities for effective use of this technique, provided the results of an up-to-date census are then available. The method is most useful shortly after a census when people are still at the addresses where they were enumerated. The amount of population shifting is probably great enough and the process of migration is selective enough to restrict seriously the usefulness of this approach when the latest census schedules are four or five years old.

USE OF SMALL AREA DATA FOR SAMPLING

The 1940 Censuses of Population and Housing have been so tabulated as to provide an unprecedented amount of material for relatively small areas. Because of the usefulness of such data for many purposes, particular emphasis was placed upon small areas in the housing census. Data were published by blocks for 191 cities with a 1930 population of

50,000 or more. These data include the number of dwelling units and structures, dwelling units by tenure and occupancy, average rent, number of units needing major repairs or having no private bath, number occupied by nonwhites, number with more than 1½ persons per room, and year built. As an aid to the use of these block data, the more significant items are to be shown on block maps for cities with a population of 100,000 or more. The maps greatly facilitate retail site selection, delineation of sales districts, and the setting of sales quotas within a given city. They also are basic to any study of the structure of cities and the shifts in residential neighborhoods which greatly affect sales outlets. This is the first time that detailed information of this sort has been tabulated on a nation-wide basis for such small segments of cities.

Addititional data are available for census tracts in 60 large cities (including adjacent areas in some cases) which have been subdivided into tracted areas. Although census tracts are considerably larger than blocks, they are still relatively small areas, usually between 3,000 and 6,000 in population. The population items available by census tracts include sex, age, race, nativity, citizenship, country of birth, highest grade of school completed, employment status, class of worker, occupation, and industry. Data on housing for census tracts include occupancy status, tenure, value or monthly rent, type of structure, state of repair and plumbing, size of household, persons per room, radio, refrigeration equipment, and heating fuel by type of equipment.

Finally, to supplement the block data, a special summary card has been prepared for each of the 110,000 enumeration districts outside of cities of 50,000 or more. The data available on these summary cards include total population and population density, and totals, ratios, or averages pertaining to color, sex, farm residence, dwelling units, home ownership, rent, farm population, farm acreage, crop acreage, and farm value. Since the average population of an enumeration district is about 900, and most of them are between 300 and 3,000, the enumeration district data provide information on characteristics of segments of the population that are highly suitable for sample use. The areas are compact geographical units, so that samples selected within such areas can be readily visited by a single enumerator or can be easily located on census schedules. Moreover, the enumeration districts are small enough to be homogeneous with respect to most characteristics. Hence, the variance within the area is likely to be small enough to simplify the drawing of a sample from the area.

By means of the statistics available for small areas, it is possible to work out a sample for large areas which is well stratified with respect to the characteristics that have proved most useful in marketing surveys. Some work based on census tracts has already been done along this line by one of the national organizations interested in marketing.[2]

SPECIAL ASSISTANCE AVAILABLE FROM THE CENSUS STAFF

There are at least three methods whereby these small-area data may be useful to marketing organizations and other organizations interested in market surveys. First, they make possible a substantial advance in the application of sampling theory to various types of problems. Work which has already been done by the Bureau of the Census, some of which is being described elsewhere in this JOURNAL[3] and some of which will be presented later, provides the basis for conclusions that are of much value to individuals and research organizations concerned with sampling.

In the second place, the Bureau can render important assistance to marketing and other organizations by reviewing the problems involved in a particular survey and making suggestions concerning the best sampling technique as indicated by census experience and statistics. If the desired degree of precision of results can be stated in advance, it is possible to determine the necessary size of sample for the purpose and to estimate the approximate cost of the field or office work.

The Bureau can also be of assistance in connection with the evaluation of completed sample surveys. Its experience with data on the basis of blocks and enumeration districts enables it to review the adequacy of the particular sample design which has been used and suggest how the results may be made most reliable. Analysis of the sample in terms of blocks or other small areas may be useful in revealing what can be inferred from a particular sample design, whether it was efficient, whether it tended to overweight areas of a special type, etc. For certain classes of data, the representativeness of the sample can be determined by testing it against the complete returns for 1940 and observing how closely the sample approximates the universe.

Finally, it cannot be too strongly urged that persons and agencies realizing the many advantages of sampling procedures for marketing studies take the opportunity of consulting the Bureau of the Census at an early stage in their work. Service can be rendered most promptly and effectively when plans for a survey are being formulated. The Bureau is hopeful that its experience and facilities may be utilized to an increasing extent by persons and organizations responsible for setting up and operating field surveys.

---

[2] For a description of the work, see *Business Week*, May 2, 1942, pp. 48–49.

[3] Benjamin J. Tepping, William N. Hurwitz, and W. Edwards Deming, "On the Efficiency of Deep Stratification in Block Sampling," this JOURNAL, March, 1943, pp. 93–100.

ON THE EFFICIENCY OF DEEP STRATIFICATION IN BLOCK SAMPLING*

By Benjamin J. Tepping, William N. Hurwitz,
and W. Edwards Deming†
Bureau of the Census

Housing statistics from the 1940 Census are available in publications of the Bureau of the Census for every occupied block in cities of 50,000 or more inhabitants, and considerable interest has been noticed in the use of these statistics for designing sample surveys. In particular, there has been interest in the use of sample designs in which restrictions are introduced comparable to those used in experimental designs of the latin square type.[1] Such sample designs will be referred to below as utilizing *deep stratification*, and these are to be contrasted with standard stratified designs. The difference between deep and simple stratification is illustrated in the following paragraphs.

> (a) *Example of stratification with complete sampling*—In selecting a sample of 16 blocks from Wilmington, all of the blocks in Wilmington might be divided into 4 geographic groups, and each of these groups further subdivided into 4 classes on the basis of the average monthly rent for the block. One block could then be selected at random from each of these 4 rent classes. This is an example of a standard design—complete sampling, each of the 16 strata being represented in the sample.
>
> (b) *Example of deep stratification with incomplete sampling*—It is possible to go further, and separate the blocks of each of the above 4 rent classes into 4 groups on the basis of the number of dwelling units in the block, and then to separate the blocks of each of these groups into 4 groups based on the average number of persons per room. The result is a table with 256 cells (4×4×4×4). It is possible to select a sample of only 16 blocks from these 256 in such manner that in the sample there is representation from each geographic location, each rent group, each group based on the number of dwelling units, and each group based on the number of persons per room. This design is balanced with respect to each criterion of stratification considered separately, but the sampling is incomplete in that only 16 out of 256 cells within the table are represented.

The purpose of this paper is to present the results of part of a series of investigations that are being carried out on the experimental and theoretical evaluation of the sampling errors and biases of several sample designs involving deep stratification, and to compare them with some sample designs commonly used. The experiments were constructed so as to make possible an examination of changes in the efficiency of a design that result from adding controls beyond the point where each stratum can be represented in the sample, but retaining the same size of sample. The restricted designs investigated are some that have been proposed for possible use.

---

* This paper was prepared for presentation at the 104th Annual Meeting of the American Statistical Association which was to have been held at Cleveland, Ohio, December 29 to 31, 1942.

† The authors acknowledge with pleasure and gratitude many valuable suggestions from Mr. Morris H. Hansen of the Bureau of the Census, not only in the planning of the experiments, but also in the preparation of the manuscript.

[1] Adaptation of experimental designs of the latin square and graeco-latin square types to sampling with "deep stratification" was originated some years ago by Lester R. Frankel and J. Stevens Stock, and was reported in a paper "On the Sample Survey of Unemployment," this Journal, Vol. 37, March 1942, pp. 77–80.

SOME DEFINITIONS

A single sample, for our purpose here, will consist of the data obtained for some one population characteristic, from a set of blocks drawn from a city according to some rule. A sample design is characterized by the rule for drawing the blocks, and the procedure for calculating the sample results (the estimating technique). Any number computed from the characteristics of the people or dwelling units in the sample is a sample result, an example being the average rent paid per dwelling unit, calculated from all the dwelling units in the sample. Successive samples, i.e., successive sets of blocks drawn by successive applications of the rule, will generate *a population of results.* The bias of a particular sample design, for a certain population characteristic, is the difference between the average estimate obtained from successive samplings, and the value of that characteristic that would be obtained by extending the survey to the entire city. A bias will arise if

> (a) the method of selecting the sample is such that the blocks have different probabilities[2] of being drawn into the sample, and if
>
> (b) the estimating technique fails to correct for these differences in the probabilities.

The appearance of the successive results for a biased design[3] will be something like those shown in the chart (*q.v.*), if the sampling is random.

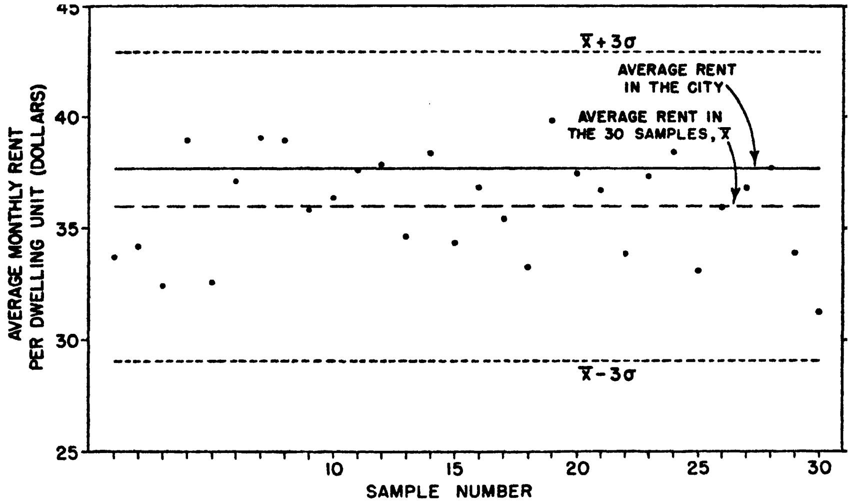
The standard error of a sample estimate is a measure of the average discrepancy between the sample estimate in a large number of successive samples, and the average of the sample estimates. It can be estimated from actual repeated samplings in an experiment, or calculated from the formula for the standard error arrived at through theoretical considerations of the sampling design, if such a formula has ever been derived. These methods can be used interchangeably if the sampling demonstrably conforms closely enough with the theory that gave rise to the theoretical standard error.

---

[2] The "probability that a block will be included in the sample" is here the ratio of the number of possible samples that contain the block to the total number of possible samples, when the combinations of blocks are made up according to a given sample design. With sufficient care and experience in carrying out an actual sample survey, it is possible not only to assimilate the theoretical calculations, but to be sure of this in advance.

[3] These types of biases are ordinarily much larger than another kind of bias that will be ignored here, specifically, a small bias that arises in estimating a ratio (e.g., average rent), when the estimates of both the numerator and the denominator are unbiased. The bias thus arising in a ratio is often introduced deliberately when it is known to be small, in order to gain the advantage of a smaller variance and the consequent greater reliability of the sample results.

RESULTS OF REPEATED SAMPLING ACCORDING TO PLAN III



NOTE: Each point shows the average monthly rent per dwelling unit, estimated from a sample of 22 blocks drawn according to Plan III. Observe the wide variation in these estimates, and the fact that their average is $1.70 less than the average monthly rent for the city.

SOME EXPERIMENTS WITH BLOCK SAMPLES FROM WILMINGTON

Each of the samples discussed in this paper consists of 22 blocks, which are approximately 2 per cent of all the blocks in Wilmington. While for most purposes this sample size is too small to give useful estimates of the characteristics of the city, it should be sufficient to distinguish between the efficiencies of the various sample designs under investigation. The standard error of a sample design is, of course, a function of the size of the sample. For the standard designs, the standard error of a sample estimate will vary exactly or approximately as the reciprocal of the square root of the size of sample; for the deep stratification designs this simple relationship does not necessarily hold. The reliability of the estimates for any of the sample designs could be improved simply by making the sample larger. Also, much better estimates of the population characteristics could be obtained with the same number of dwelling units by subsampling dwelling units within the blocks; i.e., by selecting a larger number of blocks but enumerating only some of the dwelling units in the blocks selected. It was not feasible to include subsampling in the present set of experiments. However, both of these improvements (larger sample and subsampling) can be applied equally well to the standard designs (A, B, C, D) and the deep stratification designs (I, II, III, IV). Comparisons of the biases and standard errors among all these designs are therefore useful in spite of the fact that in practice any of them might be modified.

DESCRIPTIONS OF THE DESIGNS INVESTIGATED

The sample designs that this paper deals with are described in the following outline. The four standard designs used here for comparison are denoted by A, B, C, D, and the designs involving deep stratification are denoted by roman numerals.

PLAN	STRATIFICATION	SAMPLING
A	None	22 blocks drawn at random, without replacement
B	Geographic only. The 1100 blocks of Wilmington are divided into 22 geographic groups, each containing the same number of blocks	One block drawn at random from each group
C[4]	Geographic location and average monthly rent per dwelling unit. The blocks of Wilmington are divided into 5 geographic groups, each with the same number of blocks. Within each of these groups they are divided into 4 rent classes, in such fashion that each class contains the same number of *dwelling units*, or as closely so as possible. The number of dwelling units in each of the 4 rent classes of one geographic group is not necessarily equal to the number of dwelling units in each of the 4 rent classes of some other geographic group	One block drawn at random from each location-rent class
D[4]	Geographic location and number of dwelling units per block. The blocks are divided into 5 geographic groups each with the same number of blocks, as in Plan C. Within each such group, they are divided into 4 size	One block drawn at random from each location-size class

---

[4] Note that Plans C and D, by design, provide for 20 strata, and hence for the inclusion of only 20 blocks in the sample. In comparing these plans with the other designs, their sampling errors will be adjusted to an equivalent sample size of 22 blocks.

	classes,[5] in such fashion that each class contains the same number of *dwelling units*, or as closely so as possible. The number of dwelling units in each of the 4 rent classes of one geographic group is not necessarily equal to the number of dwelling units in each of the 4 rent classes of some other geographic group	
I	The blocks of the city are divided into 22 geographic groups, each containing the same number of blocks. The blocks of a geographic group are then divided into 4 size classes,[5] each with the same number of blocks. Within each size class, the blocks are divided into 4 rent classes, each with the same number of blocks. Thus, each geographic group is broken into 16 cells (4×4), and every cell contains the same number of blocks	One block drawn from each geographic group, with the restriction that in any set of 4 consecutive geographic groups, no two blocks are to be drawn from the same rent quartile or from the same size quartile
II	Same as Plan I, except that the rent classes are determined by dividing the blocks in an entire geographic group into 4 rent classes, each with the same number of blocks. The 16 size-rent classes no longer necessarily contain the same number of blocks	Same as Plan I
III	Same as Plan II, except that the quartiles for rent and size are determined by the number of *dwelling units* in the quartile, as in Plans C and D, rather than by the number of blocks. The 16 size-rent classes do not necessarily contain equal numbers of dwelling units	Same as Plans I and II
IV	Same as Plan III	One block from each geographic group, with the restriction that in any set of 16 consecutive geographic groups, no two blocks are to be drawn from the same size-rent class

In the case of all the designs except Plan II, the estimate of the proportion of dwelling units with a given characteristic is made by dividing the number of dwelling units with the characteristic in the sample by the total number of dwelling units in the sample. For Plan II, the estimates are made by weighting the number of dwelling units in each sample block by the number of blocks in the stratum from which the sample block was selected.

---

[5] The size of a block is the number of dwelling units that it contains.

RESULTS AND REMARKS

The table shows the results that were obtained from the four designs involving deep stratification, and the results that would be obtained, on the average, from the four "standard" designs. For Plans A, B, C, and D, the standard errors and biases were calculated from appropriate approximate formulas, while for Plans I, II, III, and IV, they were estimated from repeated sampling. Plan A is unbiased, like Plans I and B, because the sampling was carried out so that the probability of being drawn into the sample is the same for every block.[3] Plan II is also unbiased because, although the probabilities there are different, the sample results were computed in a manner that takes these differences into account and adjusts for them (*vide supra*). Plans C and D, and III and IV, on the other hand, are biased, the bias being almost as great as the sampling error itself for some of the plans and for some of the characteristics estimated.

It sometimes happens that both the bias and the variance of a biased estimating technique are so small that the biased technique is preferable to an unbiased technique having a larger variance. An example was mentioned in footnote 3. In comparing the utilities of various designs, the losses and gains of both biases and variances should always be taken into account. The purpose of presenting these results for Wilmington is to furnish some evidence to assist decisions with regard to comparisons between biases and variances.

The table shows that every design except Plan II yields a gain in information over unrestricted random sampling (Plan A). With the standard stratification designs B, C, and D, it would be most unusual, although possible in sampling from finite populations, to sustain losses over random sampling, and the standard designs actually used did lead to gains over Plan A as expected.

Plan I is the same as Plan B except for the additional controls in the former furnished by size and rent. A comparison of the results indicates that the additional controls used in Plan I may produce a gain in information for some characteristics, but a loss of information for other characteristics.

Plan III is a restricted design that appears to give a better allocation of blocks than Plan I does to the strata in which there may be a greater variation of the characteristics in question. While it is not clear why this design should be more efficient than Plan I, repeated samplings of Wilmington indicated that for the proportions owner-occupied, and needing major repairs or with no private bath, Plan III gives estimates that are closer to the values for the city than Plan I gives, since in spite of the fact that the sampling bias in Plan III is of a magnitude comparable with the sampling error, the variance is small. This gain is sufficiently striking in Wilmington to justify the research that is now being conducted to discover how the efficiency of this design is affected by such things as the concentration or dispersion of a population characteristic (e.g., vacancy) over the area of the city being sampled.

Plan IV involves a simple modification of the sampling restriction used in Plan III. This modification has the effect of requiring not only that each rent quartile and each size quartile shall be represented in the sample, but that every possible combination of rent and size shall be represented. It appeared, however, that this design is less efficient than Plan III, and that there is even a loss compared with Plan B.

A more efficient design[6] than those described is one in which the blocks are stratified by size in such manner that each of the cells contains the same number of dwelling units. The sampling is to be carried out so that the probability[2] that a block in a given cell be included in the sample is in proportion to the size of the block,[5] but a constant number of dwelling units is to be drawn into the sample from each block, regardless of size. This design is unbiased, and has a smaller variance than any of the four Plans C, D, III, or IV. This scheme of sampling applied to the cells of Plans C, D, III, and IV, even as they are, would reduce their biases and variances, in spite of the fact that in these plans the cells do not necessarily contain the same number of dwelling units. This reduction in bias should arise because the expected values of the estimates are essentially the averages of the block averages, while for the sampling as proposed by Hansen the expected values of the estimates are the averages of the cell averages. If the number of dwelling units per block were the same for all blocks, the average of the block averages would be equal to the average of the cell averages, and both would be equal to the average for the city. However, when the number of dwelling units per block varies from block to block, the average of the cell averages is closer than the average of the block averages to the average for the whole city, except in rare circumstances. Experimental work with this design has not thus far been feasible.

While most of the deep stratification designs were substantially better than unrestricted random sampling in the Wilmington experiment, the fact that Plan II led to larger errors indicates that in some circumstances, deep stratification will lead to a substantial loss over random

---

[6] The principle of reducing the variance and the bias by making the probability that a block be included in the sample proportionate to its size, and then subsampling a fixed number of dwelling units from each block, is an innovation due to Morris H. Hansen.

STANDARD ERRORS AND BIASES, FOR SELECTED CHARACTERISTICS

The mean square errors for Plans I, II, III, and IV are estimated from the repeated sampling, whereas the mean square errors for Plans A, B, C, and D are calculated

Sample design	Characteristic		
	Proportion owner-occupied	Proportion needing major repairs or with no private bath	Proportion with 1.51 or more persons per room
Proportion for the entire city	.383	.258	.031
Plan A (unbiased)			
Standard error	.055	.066	.012
Plan B (unbiased)			
Standard error	.046	.057	.011
Plan C			
Bias	−.008	.001	−.000
Standard error	.046	.052	.010
Plan D			
Bias	−.027	.000	.002
Standard error	.041	.055	.011
Plan I (unbiased)			
Standard error	.048	.055	.008
Plan II (unbiased)			
Standard error	.062	.070	.008
Plan III			
Bias	−.027	.021	.005
Standard error	.031	.039	.010
Plan IV			
Bias	−.027	.003	.005
Standard error	.040	.048	.010

sampling. Deep sampling is thus not something to be introduced indiscriminately, though under some circumstances, deep stratification may indeed have greater efficiency than the standard designs. Of course, the experiments described in this paper are limited to block sampling in Wilmington and are not conclusive; further research on the subject is needed. Theoretical investigations of sampling errors and experimental work to determine the contributions to sampling error that may arise in areas of various types and for various characteristics are now being carried on in the Bureau of the Census.

THE INTEREST RATE AND SAVING*

By Donald S. Tucker
Massachusetts Institute of Technology

Saving has been represented frequently as elastic with respect to interest so that an increase of the interest rate would cause a relatively large increase in the volume of saving. The year 1932 gave an exceptional opportunity for testing the accuracy of this view of the interest rate as an inducement to save. The premium for liquidity in 1932 was high; but the long-term interest rate was also high. The premium on liquidity caused safe short-term and other highly liquid bonds to sell at high prices and at low yields. Less liquid securities sold at prices so low that the yield on long-term investments was clearly well above normal. Did this call forth an increase of saving? A survey was projected.

Preliminary discussion with various savers made it obvious that an inquiry on this subject could not be handled through a questionnaire. (1) Diminution of the family income had eliminated saving from many budgets. (2) Fear of possible further diminution or even of complete cessation of future income was causing many families to struggle frantically to curtail their current expenditures. (3) The interest rates paid by savings banks, the dividends paid by stocks and the coupon rates paid by bonds had all been either stationary or declining. It was necessary in almost every instance to explain that lower prices for fixed-income securities meant a higher current yield and a higher rate of interest.

Among about two hundred investors who were questioned during this preliminary investigation, only two were found who were both trying to save a larger percentage of their incomes in 1932 than in 1929 and thought that they were influenced in this by the interest rate rather than by a fear of further diminution of their incomes. In both instances the increase in the percentage saved was small. Because both of these savers had suffered from the general diminution of current income, absolute savings were in both instances smaller than in the previous years of prosperity. These two families were, however, trying to respond to the inducement of a higher interest rate. These were the only two out of about two hundred questioned.

The amazing thing about this investigation, however, was not the small number of those affected by the interest rate; it was an even more fundamental fact. Economists, statisticians, brokers and investment bankers, in other words, those professionally interested in the financial markets recognized readily, it must be admitted, that low prices for fixed income securities meant higher interest rates; but others, even highly successful investors, had not thought of it. All those questioned understood this relation readily when it was explained—perhaps because this small investigation was limited to a few intelligent and highly successful groups. Many of those questioned had been familiar with this relation for many years and some even were still receiving financial "services" which reported bond yields regularly. In the excitement of 1932, however, this portion of these reports seems to have been neglected. In every instance those questioned agreed after some discussion that current security prices then indicated a high interest rate, a high premium for risk and a high premium for liquidity. The high interest rate had simply no special significance for them as individuals at that time. They had not even bothered to estimate the rate then current.

---

* This paper was prepared for presentation at the 104th Annual Meeting of the American Statistical Association which was to have been held at Cleveland, Ohio, December 29 to 31, 1942.

If in prosperous and successful groups less than ten per cent had recognized the existence of high interest rates, the effect of such interest rates on the behavior of the general population must be negligible. A large scale test of this conclusion was, however, impossible. It was necessary to distinguish between a shift of the curve and the elasticity of a supply curve. Savers would discriminate between them only after questioning and discussion. Because employment of a staff was impossible, the survey was discontinued. This abortive study failed therefore to prove anything. It was, however, extremely suggestive.

Low interest rates in more recent years have apparently done very little to discourage financial saving. That interest rates affect the willingness to invest is not questioned. The demand for investment funds and the supply of capital money available for investment are both affected apparently by the interest rate; but financial saving is only one of the streams that feed the reservoir of capital money (or uninvested funds) and current investment is only one of the outlets from that reservoir. This survey made no effort to measure the outlets. It made no effort to measure real liquidation, the other great inlet to this reservoir. Even with respect to financial saving it proved nothing because of its small size; but it suggested strongly that the amount of income saved is not affected perceptibly by short-period changes in the rate of interest. In no single instance among those questioned did the 1931–1932 rise in long-term interest rates result in an actual increase in the amount of money saved.

NOTE ON DR. BERKSON'S CRITICISM OF TESTS OF SIGNIFICANCE

By R. A. Fisher
The Galton Laboratory

When, about eighteen years ago, I was looking for a good example to illustrate the test of straightness appropriate to regression lines, I felt myself particularly lucky to find one in which the departure from linearity was of a somewhat unusual kind, and in which a superficial examination by graphical methods, without submitting the subjective impression to any objective test, was likely to be misleading. The case chosen was from a comprehensive paper by A. H. Hersh on the influence of temperature on the number of eye facets developed in *Drosophila melanogaster* in a number of homozygous and heterozygous phases of the bar factor. Several of these phases showed to graphical inspection a remarkable discontinuity in the rate or direction of change in the neighborhood of 24° C. The heterozygote between the wild gene and that known as ultra-bar, on the other hand, gave frequencies which, if examined uncritically, or by methods capable of detecting only the obvious, would have failed to indicate that anything remarkable was happening at the critical temperature. It was, therefore, of some little interest to see whether a new criterion based on the rigorous solution of a problem of distribution then but recently cleared up, would enable the experimenter with such material to detect features of importance in his data, which, without such aid, would have escaped his notice.

Dr. Joseph Berkson writing, curiously enough, in 1942, when the advantages of objective tests are at least more widely appreciated than they were, refers to this example in the September number of this Journal. He has drawn the graph. He has applied his statistical insight and his biological experience to its interpretation. He enunciates his conclusion that "on inspection it appears as straight a line as one can expect to find in biological material." The fact that an objective test had demonstrated that the departure from linearity was most decidedly significant is, in view of the confidence which Dr. Berkson places upon subjective impressions, taken to be evidence that the test of significance was misleading, and therefore worthless.

It is not my purpose to make Dr. Berkson seem ridiculous, nor, of course, to prevent him from providing innocent amusement. Had he looked up Hersh's original paper he would have been spared a blunder, but we should have lost an example of the dangers of authoritarian judgment, based on subjective impressions, which even at the present date may be of value. Evidently, good biological data, examined by accurate methods, are capable of being much more informative than Dr. Berkson imagines. It is very well worth while to be reminded that general condemnations of "biological material" based on limited experience, as Dr. Berkson's judgment must be, may vastly underestimate the cogency of the evidence which careful and extensive work neatly provides.

A further confusion arises in Dr. Berkson's final comment on this example: "In this case my own judgment would be, not that the regression is non-linear, but that the temperature has varied during each or some of the experiments. At least that would explain the small *P*." As experimenters well know, one of the commonest uses for tests of significance is to detect errors of technique, or to confirm the belief that the technique has been adequate. Improbable as such an explanation seems in the case of Hersh's data, it is in general true, and important, that a discrepancy between observation and expectation may be due in fact principally to imperfections in the actual conduct of the work. If, for example, it could be supposed that the extensive series of observations at temperatures 23° and 25° C. had in error been completely interchanged, the evidence for non-linearity in the case of this genotype, though not of the others, would disappear. The fact that the test of significance has shown the series of facet numbers to be non-linearly related to the series of temperatures *recorded* would, if there really were reason to expect them to be linear on the true temperatures, have aided in the detection of such an error. What Dr. Berkson fails to realize is that the judgment, from inspection, that the line appears as straight as one ought to expect, would have given no aid whatever towards discovering the cause of any real anomaly, whatever the cause might be, because that judgment in effect denies the evidence that any real anomaly exists.

A. MANUEL FOX, 1889–1942

ABRAHAM MANUEL FOX, for many years a member of the American Statistical Association, died at Chungking, China, on June 21, 1942, at the age of 53. At the time of his death he was engaged in his duties as a member of the American-Chinese Stabilization Fund Board.

Mr. Fox was born in Philadelphia. He studied civil engineering and economics at Cornell University, the University of Michigan, and New York University. From the latter he received the degree of Master of Arts in 1924. Before entering government service, Mr. Fox for several years was engaged in expert economic research for the State of Michigan and for leading private corporations, and in teaching economic subjects in the College of the City of New York.

For many years Mr. Fox was active in the work of the United States Tariff Commission. He joined the staff in 1923 and soon thereafter became chief of the Economics Division. In 1934 the position of Director of Research was created and he was appointed to fill it. He became a member of the Tariff Commission in 1937, resigning in 1941 to join the American-Chinese Stabilization Fund Board.

During his many years as a member of the staff of the Tariff Commission Mr. Fox was the recognized leader of its research activity. He was largely responsible for the thoroughness and soundness of the wide variety of investigations conducted by the Commission. One of his most important contributions was the planning and directing of the extensive report entitled *Economic Analysis of Foreign Trade of the United States in Relation to the Tariff*, prepared in response to a Senate Resolution and published in 1933.

Both during his incumbency as Director of Research and during his membership on the Tariff Commission, Mr. Fox participated extensively in the government's trade-agreements program. He directed the preparation of a large mass of information furnished to the Congress in connection with the discussion of the enactment and subsequent extensions of the Trade Agreements Act. He had charge of the preparation, by the commodity experts and economists of the Tariff Commission, of digests of information concerning every commodity on which possible concessions by the United States were to be considered by the interdepartmental Trade Agreements Committee. He was himself a member of that committee and exercised much influence in its decisions. He was also for some time Vice Chairman of the Committee on Reciprocity Information, whose function is to receive written and oral statements from interested parties concerning possible concessions by the United States in trade agreements, as well as concessions to be sought from the foreign countries. For a number of months Mr. Fox was the head of an economic mission which visited Venezuela, at the request of the government of that country, and made important recommendations regarding tariff and other economic matters.

The American-Chinese Stabilization Fund Board consisted of five members—an American, an Englishman, and three Chinese. Mr. Fox's activities as a member were even more varied and responsible than the name of the board would suggest. In a broad sense he was a general economic adviser to the Chinese Government and a general adviser to the United States Government with respect to all economic relations with China. For several months not long before his death he was in Washington participating in important financial negotiations between the two countries. During most of the time he was in China, Mr. Fox was either in Hong Kong, from which he and his associates had an exciting escape, or in Chungking.

The following is an extract from a resolution by the United States Tariff Commission concerning the death of Mr. Fox:

> Throughout his nearly two decades of service with the Tariff Commission Mr. Fox showed unusual energy; remarkable thoroughness, keenness and ingenuity in the conduct of research; and sound judgment on economic problems and on public policy. In his personal relationships he combined inspiring leadership with friendliness and justice. He sought always the public interest. The qualities which Mr. Fox had manifested in his service with the Commission were equally conspicuous in his service in China, which, moreover, involved often great personal hardship and danger. By his sudden death in the full vigor of his powers, the United States, the Republic of China, and the other United Nations have lost a highly efficient and deeply patriotic public servant.

E. DANA DURAND

United States Tariff Commission

DAVIS RICH DEWEY, 1858–1942

Davis Rich Dewey who died in Cambridge on December 13, 1942, was secretary of the American Statistical Association for twenty years (1886–1906). He was also throughout that period editor of its *Publications*, though during the latter portion of this period he was assisted by a committee. Because these Publications developed later into the *Journal of the American Statistical Association*, he might almost be considered the first editor of this Journal. He was certainly the Statistical Association's first editor. In addition to being present at the founding of the American Economic Association in 1885, he became its vice-president in 1908 and its president in 1909, he was also an early member of its publications committee and became finally the first editor of the *American Economic Review*. He served as managing editor of that magazine for thirty years (1911–1940). As editor, as statistician and as economist, Professor Dewey was interested always in the accurate reporting of fact.

He was born at Burlington, Vermont, in 1858, and graduated from the University of Vermont in 1879, taking Phi Beta Kappa honors. He taught for one year in a district school at North Underhill, Vermont, and in the following year was principal of a country academy at Underhill. In 1881 he was appointed principal and teacher of Latin and Greek at the Hyde Park High School, near Chicago. In 1883 he entered the graduate department of Johns Hokpins University, secured a fellowship there, and in 1886 received his doctor's degree in economics. While at Johns Hopkins he connected himself with Bradstreet's and was a regular correspondent of the publication; he spent the summer of 1885 in editorial work in the offices of Bradstreet's.

Coming directly from Johns Hopkins in 1886 to an instructorship in history and economics at the Massachusetts Institute of Technology, he was made professor of economics and statistics in 1893. He was head of the department of economics and statistics until his retirement, and was in charge of the course in engineering administration from its beginning in 1913 until 1931, when a separate department of engineering administration was established.

Always interested in questions of public welfare, Professor Dewey served on many state and national commissions. In 1894 he was chairman of the Special Commission in Massachusetts to Investigate the Subject of the Unemployed; in 1896 he was a member of the Board to Investigate Charitable and Reformatory Institutions of Massachusetts; in 1903 he was a member of the Commission on Relations between

Employer and Employee. As a result of these investigations he shared in the publication of several reports relating to the respective subjects. From 1897 to 1905 he was a member of the Municipal Department of Statistics of the City of Boston. In 1901 he was appointed special expert agent on wages of the Twelfth United States Census; in this latter capacity he conducted the research necessary for the publication in 1903 of one of the most significant contributions of the Census Bureau to wage statistics, a volume entitled *Employees and Wages*, a collection of wage rates covering thirty-four industries. Between 1928 and 1932 he served on three emergency boards and on one board of arbitration appointed by the President of the United States to investigate railroad disputes. These cases involved several railroads of the Middle West and some nation-wide labor organizations: the Kansas City, Mexico and Orient system; the Western Railways *versus* the Order of Railway Conductors and Brotherhood of Railway Trainmen; the Illinois Central Railroad, the Yazoo and Misssissippi Valley Railroad and the Gulf and Ship Island Railroad; the Louisiana, Arkansas and Texas Railway Companies.

He was connected with educational institutions in various capacities. He received an honorary degree of LL.D. from the University of Vermont in 1910. He was a trustee of Massachusetts State College from 1909 to 1939. At various times he gave courses in economics at Johns Hopkins University, Wellesley College, Harvard University, Columbia University and the University of California. His interest extended also to secondary education; and he became recently a trustee of the Browne and Nichols School for Boys in Cambridge.

Among Professor Dewey's more important publications are *Financial History of the United States*, in recognition of which he was awarded in 1904 the John Marshall Prize by Johns Hopkins University and which went into its thirteenth edition in 1941; he was editor of the two-volume work, *Discussions in Economics and Statistics* by Francis A. Walker (1899); articles in *Literature of American History* edited by Larned (1902); *National Problems, 1885–1897* (1907); *State Banking before the Civil War* (1910); joint author with Mr. M. J. Shugrue of *Banking and Credit* (1922); articles for Palgrave's *Dictionary of Political Economy;* the *New International Encyclopaedia Americana; Encyclopaedia Britannica; American Year Book*, and the *Commonwealth History of Massachusetts.*

His friends loved to think of him surrounded by his family in his summer home at New London, New Hampshire. The flower gardens, which he tended himself, were notable and the view from his farmhouse was picturesque. Occasionally Professor and Mrs. Dewey left this delightful location for a summer of travel—in 1923 to the Pacific Coast and in 1929 to Europe when he attended the annual meeting at Warsaw of the Institut International de Statistique; in 1936 to Mexico for another meeting of the I. I. de S.

In 1886 Dr. Dewey was married to Mary C. Hopkins of Madison, Wisconsin, who survives him. Their children are Mrs. A. Barr Comstock, now living in Washington, and Colonel Bradley Dewey, president of the Dewey and Almy Chemical Company and now deputy rubber administrator. He is also survived by his brother, Dr. John Dewey, by six grandchildren, Dr. Bradley Dewey, Jr., Dr. Davis R. Dewey, II, Marguerite Mellen Dewey, Ann Dewey, Ensign A. Barr Comstock, Jr., Dorothy Comstock and by a great grandchild, Margot Dewey.

No man ever remained long in contact with him without being helped by his kindness and inspired by the nobility of his character.

BEATRICE A. ROGERS and
DONALD S. TUCKER

CHARLES P. NEILL, 1865–1942

THE PASSING of Dr. Charles P. Neill on October 3, 1942, at his home in Washington, D. C., after a protracted illness, closed a long and eminent career. Among the positions he occupied in a life crowded with notable accomplishment were: Professor of political economy at the Catholic University of America; Assistant Recorder of the Anthracite Coal Strike Commission; United States Commissioner of Labor; Member of the United States Immigration Commission (1907–10); Manager of the Bureau of Information of the Southeastern Railways (1915–39); Member of the United States Coal Commission (1922–23); and Umpire for the Anthracite Conciliation Board for over twenty years.

Dr. Neill, a son of Julia and James Neill, was born in Rock Island, Illinois, in 1865, and his early youth was spent in Austin, Texas. He attended the University of Texas, received his Bachelor's degree from Georgetown University (1891), his Master's degree from the University of Notre Dame, and his Doctor's degree from Johns Hopkins University (1897). While still pursuing his studies at the last mentioned institution he became instructor at the Catholic University and was its professor of political economy when he resigned in 1905 to become United States Commissioner of Labor.

Notre Dame University conferred upon him the Honorary Degree of Doctor of Laws in 1908, and in 1922 awarded him the Laetare Medal for his outstanding services in controlling child labor in the South and for the mediation of labor disputes.

As Commissioner of Labor under Presidents Theodore Roosevelt and Taft, Dr. Neill's services were regarded as so valuable that he was reappointed at the beginning of President Wilson's administration.

During the great part of Dr. Neill's incumbency, 1905–1913, the Bureau of Labor was in the Department of Commerce and Labor and many important functions now performed by various other Federal agencies devolved upon that Bureau which was a pioneer in the promotion of labor legislation. It may be recalled that workmen's compensation legislation, which is now taken for granted, was then considered as a remarkable step forward in social progress in the United States.

The Bureau was also charged with the investigation of the conditions of woman and child labor, and under the direction of Dr. Neill made a most comprehensive study of this subject, the results of which were published in 19 volumes.

Under the Erdman Act, which constituted the embryo of later elaborate railroad legislation for the preservation of industrial peace, the Chairman of the Interstate Commerce Commission and the United

States Commissioner of Labor were constantly called upon in person to use their good offices in railway labor controversies. Judge Martin A. Knapp, then chairman of the Interstate Commerce Commission, and Dr. Neill were instrumental in averting numerous important strikes, some of which would have been colossal in their scope. The prevention of a threatened national tie-up of telegraph lines was among the triumphs for the Bureau of Labor in the field of industrial relations.

In 1906 President Roosevelt appointed Dr. Neill and James B. Reynolds to investigate the meat-packing industry in Chicago. The findings were embodied in a courageous report which aroused nation-wide interest.

While he was Commissioner of Labor, Dr. Neill was also a member of the United States Immigration Commission (1907–10), which made a study of the immigration problem from innumerable angles and issued its findings in 41 volumes.

Dr. Neill served as Umpire for the Anthracite Conciliation Board from 1906 to his resignation in 1929, combining this arduous office with various other important activities. The following excerpts from an editorial in *The Scranton Times*, of October 5, 1942, headed "The Anthracite Law Giver," summarize his work in this field.

> As long as anthracite coal is mined the generations of mine workers . . . will be indebted to Dr. Neill.
>
> The interpretations or decisions made by Dr. Neill still represent the body of the law of the anthracite. Except for modifications due to the wider introduction of machinery and the development of mine engineering and new work, the law of the anthracite industry established by Dr. Neill is the rule today as it affects wages and working conditions directly and human relations potently in a great and unique industry.

Dr. Neill was long a Fellow of the American Statistical Association and in 1916 he served as our President.

For many years Dr. Neill was prominent in the national Capital's civic and welfare organizations, served on the Community Chest, and was for eight years Vice President of the District of Columbia Board of Charities.

Dr. Neill was a staunch Catholic and a dominant influence in National Catholic movements in the field of social action. He gave generously of his time on national committees. He was deeply interested in the promotion of the National Catholic Welfare Conference, especially in its formative years, and was the first director of the National Catholic School of Social Service. After his resignation from that office he continued to lecture on economics at that institution.

A. F. HINRICHS

Acting Commissioner of Labor Statistics

NORMAN JOHN SILBERLING, 1892–1942

NORMAN J. SILBERLING died on October 18, 1942, at his residence in Palo Alto, California. At the time of his death he was Vice-President of the American Statistical Association, covering the field of facts and methods bearing upon economics and economic theory. The Association has lost a valuable member.

Dr. Silberling was born in Cleveland, Ohio, in 1892. In his undergraduate days at Harvard he developed a keen interest in economics. After receiving the Bachelor's degree, *magna cum laude*, in 1914, he continued his economic studies as a graduate student, receiving the Master of Arts degree in 1915 and Doctor of Philosophy in 1919. His doctoral dissertation was on the *History of Theories of Money and Credit from Adam Smith to John Stuart Mill.* It pointed the way toward his life work. The following year, Dr. Silberling was appointed instructor of the Department of Economics at Harvard, and during 1920–1921 held an assistant professorship at Dartmouth College.

In the fall of 1921 he went to the University of California as associate professor of Economics, receiving his full professorship in 1929. In 1930, his status was changed to that of part-time lecturer in order that he might give more of his time to the research organization which he had established. In 1935 he was appointed Professor of Business Research in the Graduate School of Business. He continued in this capacity until his death.

Dr. Silberling was an able and inspiring teacher but his major interest lay in financial research. In *The Silberling Reports*, later *Graphic Survey of Industry and Finance*, and in confidential weekly letters, circulated among his business and investment clients, he made interesting statistical analyses of the economic development of the western region which should be of real value to economic historians of the interwar period. His greatest satisfaction came from his labors in what he called the sphere of business dynamics. In this field his book, which fortunately he had completed with the exception of minor revisions in the later chapters, is a major contribution.

REPORTS OF THE ASSOCIATION

Report of the Board of Directors

As nineteen hundred and forty-two, one of the most momentous years in the history of the nation and the world, draws to a close, we again undertake to review the activities and accomplishments of our Association and to appraise the outlook for the immediate future. As in the life of the nation the year was one of rapid change and repeated readjustment to new conditions, so in the history of the Association it will be recorded as one in which precedents were superseded and traditions broken. For both patriotic and practical reasons, the Annual Meeting which, more than any other single activity, symbolizes the Association's life and progress was cancelled. To the Vice-Presidents and other officers and members who participated in the planning of the Cleveland meeting, we wish to express hearty appreciation of the efforts expended and sincere regret that the sessions as planned could not materialize. We hope that a considerable number of the papers intended for the Annual Meeting will become available to the members in the columns of the JOURNAL.

The contribution of the Association to the prosecution of the war may be measured in terms of the active participation of its officers and members in the vital work of the government war agencies, institutions of learning, and industrial concerns. Many of our members occupy key posts in the statistical branches of the government service. Others are serving as consultants on a part-time basis. College and university members are carrying heavy teaching loads and, in addition, are collaborating on important war research projects. Other members are contributing to the progress of the war through the development of new techniques, and the application of statistical controls, in industry. Your officers and Directors have devoted considerable attention to the study of ways in which the Association might contribute directly to the furtherance of the war program. Under instructions from the Board of Directors, the Secretary early in the year interviewed government statisticians, business officials and members of university and foundation staffs to obtain suggestions as to possible contributions the Association might make to the war effort. One of the Directors, serving as a committee of one appointed by the Board, has devoted much time to exploration of means by which the Association might help to increase the extent and improve the quality of applications of statistics to war problems.

Chapters: The Chapters have experienced the impact of the war in varying degree. All have been affected to some extent by the loss of active members to the armed forces, by the emigration of others from their home communities to accept positions elsewhere and by the preoccupation of the remaining members with war work. It is encouraging to note, however, that, despite these handicaps, unusually fine programs have been developed by some of the Chapters. The Washington Statistical Society and the New York

District Chapter have made tentative plans for the presentation of a selection of the Annual Meeting papers at Chapter meetings. New Chapters established during the year were the Asociacion Cubana de Estadistica Chapter in Havana, Cuba, and the Central New Jersey Chapter. The Charter of the Lehigh Valley Statistical Society, which had been inactive for several years, was revoked. The Detroit District Chapter, also inactive in recent years, has resumed regular meetings and it is anticipated that a reorganization of the Madison, Wisconsin, Chapter will be effected early in 1943. With the cancellation of the Annual Meeting in 1942 and the prospect that the usual type of general meeting will have to be discontinued for the duration of the war, the contribution of the Association to its members becomes increasingly dependent upon the vitality of the Chapters.

The Biometrics Section: During the year the Biometrics Section held joint meetings with the American Society of Biological Chemists, the American Association of Cereal Chemists, the Soil Science Society of America and the American Society of Agronomy. Additional joint sessions with five other biological societies and with the Institute of Mathematical Statistics were planned for the Annual Meeting of our Association in Cleveland and the meetings of the American Association for the Advancement of Science in New York but had to be cancelled. While the organization of additional sections in other fields and the better coordination of sections with the parent Association are included in the long range aims of the Association, the Directors feel that efforts in these directions should await a more propitious time than the present.

Committees: In spite of the handicaps imposed by the national emergency, most of the Association's committees continued to render effective service in their respective fields. The Census Advisory Committee held three two-day meetings. Consideration was given to the need for expanding the staff of the Census Bureau and retaining a nucleus of well-trained professional persons for carrying on current work, conducting intercensal projects and planning for the decennial censuses. There was also extended discussion of questions relating to the establishment of a national population registration system and the role of the Bureau in the prosecution of the war. The Committee on Census Enumeration Areas was effective in promoting the distribution of the 1940 census tract publications of the Bureau of the Census through extensive correspondence with representatives of the various census tract cities. The Joint Committee on Occupational Classification served as a clearing house for occupational classification questions and advised various agencies undertaking occupational coding. A subcommittee was actively engaged throughout the year in checking the convertibility of the titles in the Supplement to the United States Employment Service *Occupational Dictionary* to the Convertibility List and in reviewing the allocation of Army occupations in the *Occupational Dictionary.* The Joint Committee on Relief Statistics continued to function as an advisory committee in connection with various aspects of the work of the Social Security Board and individual members of the committee were called upon for advice at various times.

During the year an advisory committee to the Federal Security Administrator in the field of morbidity statistics was appointed to serve under the chairmanship of Dr. Lowell J. Reed. The Committee on Hospital Statistics was discontinued. To meet the needs formerly served by the latter Committee, Dr. Halbert L. Dunn and Dr. Edwin L. Crosby were appointed consultants from the American Statistical Association to the Committee on Accounting and Statistics of the American Hospital Association.

Members of the Joint Committee for the Development of Statistical Applications in Engineering and Manufacturing advised and consulted with government and industrial groups on the application of statistical control techniques to problems arising in connection with the war. They also cooperated in the arrangement of scientific meetings and in the organization of statistical effort in engineering and scientific fields. A subcommittee on Engineering and Scientific Graphs of the Sectional Committee on Standards for Graphic Presentation submitted to the American Standards Association its proposed standard "Engineering and Scientific Graphs for Publication." Through its representatives, the Association also participated in the work of the Social Science Research Council, the National Bureau of Economic Research, the American Association for the Advancement of Science, the Advisory Board of the *American Year Book*, the Advisory Committee to the Census Library Project and the American Documentation Institute.

Publications: Volume 37 of the JOURNAL of the American Statistical Association, published in 1942, was similar in size and content to other recent volumes. As in other years, a considerable proportion of the first two numbers in the volume was devoted to Annual Meeting papers. Volume 2 of the American Statistical Association *Bulletin* was completed. For reasons of economy, the six issues published in 1942 were much smaller than those of the preceding year.

Membership and Finances: The membership records for 1942 are in striking contrast with those for the three preceding years and reflect in the main the adverse conditions of the present war period. The number of names removed from the rolls by reason of resignation, death, or termination for failure to pay dues, totaling 400, was more than one and one-half times as large as the total number of new members and reinstatements. This resulted in a net decrease of about 200 in the membership of the Association.[1] Of those who resigned or were delinquent in payment of dues, slightly less than two-thirds joined the Association during the last three years. The number of new members enrolled during the year (232) was the same as in the year preceding that of the Centenary Celebration.

Income from subscriptions, as well as from dues, showed a substantial decrease in 1942. Nevertheless, as the result of increased income from other sources, the institution of economies in office and other expense, and the continued generosity of the American University in permitting us to retain our present quarters, the Association was able to close the year with a slight excess of receipts over expenditures. As shown in the financial statement, the Association's income for the year was $19,633.08 while the total expenses were $19,271.60. No funds were withdrawn from the Centenary Sustaining Fund during the year. The balance in this Fund at the close of 1942 was $4,802.97.

---

[1] These estimates were superseded by figures compiled at the end of the year. See the Report of the Secretary, immediately following.

Outlook for the Future: As suggested earlier in this report, your Directors believe that the present times are unfavorable to the maintenance of a program of continued expansion such as was contemplated at the time the Centenary Sustaining Fund was established. It seems to them the part of wisdom, therefore, to adopt a program of retrenchment for the duration of the war, curtailing expenses to keep, if possible, within the income available from regular sources alone, and husbanding the resources of the Centenary Sustaining Fund until such time as conditions offer greater promise of the success of undertakings which would necessitate drawing upon that Fund. When opportunities for useful service become apparent and the prospects for the successful realization of our objectives become brighter than at present, we shall feel justified in again drawing upon the Centenary Sustaining Fund to defray the expenses of an enlarged program of activity.

W. LEONARD CRUM
R. L. FUNKHOUSER
F. LESLIE HAYFORD
ALFRED J. LOTKA
LOWELL J. REED
WINFIELD W. RIEFLER
WALTER A. SHEWHART
FREDERICK F. STEPHAN

Report of the Secretary

It is customary for the Secretary to present certain statistics about the membership of the Association at the close of each calendar year. In 1942 the total membership decreased from 3,181 to 3,037, a net decline of 144 members. During the year, 232 new members were elected and 24 former members were reinstated. Two hundred and ten members were removed from the active rolls for failure to pay their dues. One hundred and seventy-five members resigned and 15 were lost by death, including three Fellows. Of the latter, Carroll W. Doten and Charles P. Neill were Past Presidents of the Association, and Davis R. Dewey was Secretary of the Association for many years and the first Editor of the *Publications of the American Statistical Association.*

Membership statement, December 31, 1942

Honorary members	14
Corporate members	4
Fellows	109
Regular members	2,910
Total membership	3,037

One Fellow was a Contributing Member during 1942. One new Life Member was elected during the year, bringing the total number of Life Members to 37.

The death of the following members was recorded during the year: Davis R. Dewey, Carroll W. Doten, and Charles P. Neill, *Fellows;* Horace R. Coddington, John R. Dower, Elmer R. Enlow, Emily I. Farnum, A. Manuel Fox, William F. French, Constance M. Gay, James L. Gore, Clyde M. Leffler, Norman J. Silberling, Walter E. Trent, and Frank W. Wheeler, *Regular Members.*

R. L. FUNKHOUSER, *Secretary*

Report of the Committee on Fellows

During 1942 the Committee on Fellows elected ten members of the Association to Fellowship. This was in accord with earlier decision by the Association that the number of Fellows should be slowly increased toward a maximum of one hundred and fifty. Those elected this year were:

C. I. Bliss
W. Edwards Deming
Edward L. Dodd
Edward V. Huntington
Robert R. Nathan
Jerzy Neyman
Frank W. Notestein
John W. Scoville
Dorothy Swaine Thomas
Rufus W. Tucker

At the beginning of 1942 there were one hundred and two living Fellows. During the year three died.

FRANK A. ROSS, *Chairman*

Report of the Treasurer

In 1942 the income of the Association from regular sources totaled $19,633, a decline of nearly $600 from the 1941 figure. Most of the reduction was in the sale of the *Index to the Journal*, which was published early in 1941. Among the other sources of revenue, declines in income from membership dues and from subscriptions were offset substantially by increases in advertising revenue, the sale of reprints, and in JOURNAL sales. Owing to the War, a number of foreign libraries and other institutions have had to cancel their subscriptions to the JOURNAL, accounting for most of the decline in income from this source.

As noted in the *Report of the Board of Directors*, there was no transfer in 1942 from the Centenary Sustaining Fund to the general account of the Association, whereas the year before the sum of $2,500 was used to supplement the regular income. The Centenary Sustaining Fund is being preserved pending the return of an appropriate opportunity to resume an expansion in the scale of the Association's activities.

A rise in paper prices and labor costs was chiefly responsible for the increase in the cost of the JOURNAL. With this exception, practically all items of expense showed reductions from the corresponding 1941 figures. Expense for the *Bulletin* was cut approximately in half by curtailing the size of the issues. The total of salaries and wages was reduced more than $500. Through the generosity of the American University in making office space available to the Association, nearly $1,000 was saved on rent. Economies were effected in the various forms of office expense.

The increase in the cost of back numbers of the JOURNAL sold during the year was the result of the enlarged volume of such sales. The expenditure of about $300, charged to the "Survey of defense statistics" was largely for additional personnel assistance in the office in order to free the Secretary to study possible steps that the Association might take to contribute more effectively to the war effort.

Aside from the non-recurring expenses in 1941 in connection with the *Index to the Journal*, the total expenses of the Association were reduced about $2,900 in 1942. As a result of this retrenchment, a balance of nearly $400 was carried to surplus notwithstanding the drop in income.

The balance sheet and other detailed financial statements, with comparative figures for 1941, are included in the *Report of the Auditors*.

R. L. FUNKHOUSER, *Treasurer*

Report of the Auditors

To the Board of Directors of
American Statistical Association

We have examined the balance sheet of the American Statistical Association as of December 31, 1942 and the statements of income and surplus for the year then ended. Our examination was made in accordance with generally accepted auditing standards applicable in the circumstances, and included such tests of the accounting records and other supporting evidence and such other procedures as we considered necessary.

The recorded cash receipts for the year were traced to the deposits shown on the bank statements and the amounts for dues and subscriptions were tested with the membership and subscription records. The paid checks and relative vouchers were inspected in support of the cash disbursements for the year. The cash balances and the securities owned at December 31, 1942 were confirmed by inspection or by certificates obtained direct from the depositaries. We did not check the membership and subscription records in detail or make any independent verification of the inventory of old journals, the office records of which are based in part on data assembled in prior years, no recent physical inventory having been taken.

The life membership reserve is computed on the basis of the combined annuity table of mortality with assumed interest at 4% per annum and an assumed annuity of $5.00 per life member, in accordance with a resolution of the Board of Directors on March 31, 1936. The amount treated as income in each year represents the excess of the reserve at the beginning of the year plus interest for the year and new life membership receipts over the required reserve at the end of the year.

In our opinion, the accompanying balance sheet and related statements of income and surplus present fairly the position of the American Statistical Association at December 31, 1942 and the results of its operations for the year, in conformity with generally accepted accounting principles applied on a basis consistent with that of the preceding year.

PRICE, WATERHOUSE & CO.

March 13, 1943

AMERICAN STATISTICAL ASSOCIATION
BALANCE SHEETS

Assets	December 31, 1942	December 31, 1941
Cash in bank and on hand	$ 6,547.86	$ 1,753.69
Accounts receivable	262.84	325.61
Investments:		
United States Savings Bonds, Series D, at redemption value	5,320.00	5,186.00
Stocks, at cost (at market quotations $4,125 and $3,875, respectively)	5,793.50	5,793.50
Inventory of old Journals, at approximate cost	1,655.58	1,779.91
Furniture and equipment, at cost less depreciation	479.49	541.81
Deferred charges (expenses applicable to subsequent year)	255.57	150.31
	$20,314.84	$15,530.83

Liabilities		
Accounts payable	$ 1,491.30	$ 1,183.18
Special account—Allied Social Science Associations		481.81
Centenary Sustaining Fund, per statement	4,802.97	3,618.45
Life membership reserve	2,341.53	2,315.79
Deferred credits (collections applicable to subsequent year):		
Dues	4,753.67	1,267.12
Subscriptions	1,572.67	1,690.76
Contributions	41.00	
Surplus, per statement	5,311.70	4,973.72
	$20,314.84	$15,530.83

AMERICAN STATISTICAL ASSOCIATION
INCOME STATEMENTS

	Year ending December 31,	
	1942	1941
INCOME:		
Dues—current year	$13,765.50	$14,002.00
Dues—prior years	70.00	31.25
Life memberships	166.89	192.60
Subscriptions	3,277.25	3,723.25
Advertising	950.19	832.62
Reprints	404.50	174.40
Journal sales	635.75	293.07
Index to the Journal, sales	23.50	566.00
Special publications	63.76	71.07
Miscellaneous	110.95	128.25
Dividends and interest (after deducting $96.48 in 1941 and $92.63 in 1942 apportioned to life membership reserve)	164.79	208.34
	$19,633.08	$20,222.85
Appropriation from Centenary Sustaining Fund		2,500.00
The Rockefeller Foundation grant for Index to the Journal (per contra)		912.74
	$19,633.08	$23,635.59
EXPENSES:		
Journal—printing, mailing and reprints	$ 4,852.81	$ 4,722.89
Bulletin	726.95	1,441.75
Salaries and wages	11,111.07	11,682.29
Unemployment compensation tax	233.40	238.96
Rent		955.00
Office supplies, printing and mimeographing	470.11	932.76
General postage and carriage	454.40	710.04
Telephone and telegraph	141.43	155.62
Travel expense—officers	246.49	387.20
Mimeographing—committees	41.09	59.46
Storage of old Journals	72.00	72.00
Cost of old Journals, sold	204.33	50.64
Miscellaneous expense	307.58	475.35
Depreciation of furniture and equipment	84.62	95.61
Membership campaign		170.80
Survey of defense statistics	325.32	
Expense—Index to the Journal:		
Expended from Rockefeller Foundation grant, per contra		912.74
Expended from Association funds		211.35
	$19,271.60	$23,274.46
Excess of income over expenses, carried to surplus	$ 361.48	$ 361.13

AMERICAN STATISTICAL ASSOCIATION
SURPLUS STATEMENT

	Year ending December 31, 1942	1941
Balance at beginning of year	$4,973.72	$4,967.24
Add—Excess of income over expenses for the year, per income statement	361.48	361.13
	$5,335.20	$5,328.37
Deduct—Appropriation to Centenary Sustaining Fund of amount equal to net proceeds from sales of Index to the Journal	23.50	354.65
Balance at end of year	$5,311.70	$4,973.72

STATEMENT OF CENTENARY SUSTAINING FUND

		Total	Year ending December 31, 1942	Prior to 1942
Contributions and pledges (relating to the support of the activities of the Association for a period of five years from January 1, 1940 to December 31, 1944)		$10,686.02		
Less—Pledges not collected at December 31, 1942 including $1,244.90 not due at that date		1,716.25		
Amounts received from contributors		$ 8,969.77	$1,065.00	$7,904.77
Interest received on bank savings account and increase in redemption value of securities		231.58	96.02	135.56
Appropriation from surplus of amount equal to net proceeds from sales of Index to the Journal		378.15	23.50	354.65
Total receipts		$ 9,579.50	$1,184.52	$8,394.98
Less—Expenses of campaign (printing, postage and temporary assistance)		776.53		776.53
Net receipts		$ 8,802.97	$1,184.52	$7,618.45
Appropriations by the Board of Directors to the general account of the Association in accordance with the budget:				
Year 1940	$1,500.00			
Year 1941	2,500.00	4,000.00		
Balance December 31, 1942, per balance sheet		$ 4,802.97		

List of Committees and Representatives for 1942

Committee on Fellows[1]

Frank A. Ross
W. Leonard Crum
John Rice Miner
Theodore H. Brown
Lowell J. Reed

Committee on Nominations

Frederick F. Stephan, *Chairman*
Joseph S. Davis
George W. Snedecor

Committee on Investments[2]

F. Leslie Hayford, *Chairman*
George O. May
Walter W. Stewart

Biometrics Section Committee

C. I. Bliss, *Chairman*
Joseph Berkson, *Secretary*
A. E. Brandt
Alfred J. Lotka
J. Neyman
W. J. Youden
Churchill Eisenhart
George W. Snedecor
Frank W. Notestein, *ex-officio*

Census Advisory Committee

William F. Ogburn, *Chairman*
Murray R. Benedict
Paul T. Cherington
J. Frederic Dewhurst
Luther Gulick
Frederick F. Stephan
Willard L. Thorp

Committee on Census Enumeration Areas

Howard W. Green, *Chairman*
Clarence E. Batschelet
W. Thurber Fales
Ernest M. Fisher
Shirley K. Hart
Charles S. Newcomb
Vergil D. Reed
Leon E. Truesdell
Lent D. Upson

Joint Committee on Occupational Classification (With the Division of Statistical Standards, Bureau of the Budget)

Howard B. Myers, *Chairman*
Gladys L. Palmer, *Secretary*
E. Dana Durand
Meredith B. Givens
Edward D. Hollander
Merrill G. Murray
Leon E. Truesdell
Emmett H. Welch

Joint Committee on Relief Statistics (With the American Public Welfare Association)

Ralph G. Hurlin, *Chairman*
Paul Webbink, *Secretary*
Neva R. Deardorff
Emil J. Frankel
Anne E. Geddes
Howard B. Myers
Melvin W. Sneed
Gaylord A. Walker
Emmett H. Welch

---

[1] With terms expiring at the end of 1942, 1943, 1944, 1945, and 1946, respectively.

[2] Appointed in accordance with By-Law 5 to supervise the investment of the Association's surplus funds.

Committee on Chapters
Robert R. Nathan, *Chairman*
Henry B. Arthur
Edwin B. George
Lester S. Kellogg
William A. Sturm
Charles W. Vickery, Jr.
Robert R. Williams

Committee on Sampling
Frederick F. Stephan, *Chairman*
W. Edwards Deming
George W. Snedecor
Alfred N. Watson
Samuel S. Wilks
Theodore O. Yntema

Advisory Committee to the National Roster of Scientific and Specialized Personnel
Halbert L. Dunn
Walter Mitchell, Jr.
Frederick F. Stephan
Ralph W. Tyler
Samuel S. Wilks

Advisory Committee on Morbidity Statistics
Lowell J. Reed, *Chairman*
E. L. Bishop
Hugo Muench
H. S. Mustard
E. S. Rogers

Representative on the Board of Directors of the National Bureau of Economic Research
Winfield W. Riefler

Members of the Social Science Research Council[3]
Stuart A. Rice
Edwin B. Wilson
O. C. Stine

Representative on the Council of the American Association for the Advancement of Science
Walter A. Shewhart

Representatives on the Joint Committee for the Development of Statistical Applications in Engineering and Manufacturing
Walter A. Shewhart
Churchill Eisenhart

Representative on the Sectional Committee on Standards for Graphic Presentation
A. H. Richardson

Member of the American Documentation Institute
Richard L. Funkhouser

Representative on the Advisory Board of the American Year Book
Frank W. Notestein

Representative on the Advisory Committee to the Census Library Project
Richard O. Lang

---

[3] With terms expiring at the end of 1942, 1943, and 1944, respectively.

BOOK REVIEWS

GLENN E. McLAUGHLIN
Review Editor

Neue Schweizerische Sterbetafeln. Bern: Beiträge zur Schweizerischen Statistik. Heft 10. 1942. 56 pp.

Some of the principal results of this valuable study of the Central Statistical Bureau of Switzerland (Eidenössisches Statistisches Amt) may be reproduced in brief. The first chapter deals with the recent Swiss life table, 1933/37, in comparison with the former eight covering the period from 1876 to 1932. Accordingly, the life expectancy at birth was (in years):

Life table	Males	Females
1876/80	40.64	43.24
1901/10	49.25	52.15
1933/37	60.68	64.65

The last values are almost exactly the same as obtained in U.S.A. 1935, for the white population (60.7 and 64.7) and in England and Wales, 1937 (60.2 and 64.4), as quoted in the international comparison of the publication. Yet they are still somewhat below the respective values in Holland, Denmark (only for males), Australia, and New Zealand for the last life table periods (around 1935). The prolongation of life in the Swiss population during the sixty years covered amounts to exactly 20 years for a new born boy and 21.4 years for a new born girl. By far the largest part of this great average gain was due to the decrease of infant and early childhood mortality. The absolute contributions in years for the different age groups are given in a separate tabulation. This contribution was for the first year of life 8.6, for age 1–4 still 3.2 years, or together more than half of the total gain in the male sex; the corresponding figures in the female sex are 7.8 and 3.6 years. The next highest gains occur in the age group 30–39 with 1.5 and 1.7 years, respectively for the sexes. In contrast to this, the gain in further life expectancy for the age groups, 70 years and above, amounts only to 0.4 for the male and 0.8 years for the female sex. This brings out rather definitely, also in the Swiss life tables, the fact that the individual life span has not very much changed after the hazards of childhood and maturity have passed. Individual life, then, still lasts "three score and ten."

The probabilities of dying per year are given later in extenso and on this basis tables of survivors comparing the dying out of a generation at different periods of the Swiss population. The "probable life duration" as characterizing the age in which one-half of the starting generation has died out, has increased from 47.3 (life table 1876/80) to 67.5 years (life table 1933/37) in the male sex; from 51.7 to 71.5 years, respectively, in the female sex. It is throughout all periods larger than the mean duration of life at birth. The development of the most important causes of death is afterwards dealt with in summarizing tabulations. Of special interest is the construction of a life table for lung tuberculosis and for cancer. Excluding all deaths from either of these causes in the life table 1929/32, life expectancy results in the following values (years):

Age (years)	Males			Females		
	All causes of death	Without lung tuberculosis	Without cancer	All causes of death	Without lung tuberculosis	Without cancer
0	59.17	60.71	60.66	63.05	64.85	64.60
40	28.58	29.34	30.32	31.42	31.96	33.13
60	13.91	14.15	15.26	15.45	15.66	16.57

The figures demonstrate that tuberculosis, although having decreased enormously during the last decades, still in the life table of 1929/32 causes a greater loss of life for the total Swiss population than cancer.[1] However, in the middle and higher ages the further expectation of life when excluding all cancer deaths is distinctly greater than when excluding deaths from lung tuberculosis. Thus the eradication of cancer would significantly increase only further life expectancy at these ages.

Further life tables are constructed for the urban and rural population. They show that at present city dwellers, especially women, have a somewhat better life expectancy than country dwellers. The last chapter deals shortly with mortality by marital status (single, married, widowed, divorced) and confirms the fact well-known from other experience that also in the Swiss life table married people have the most favorable, the divorced ones the least favorable rates of mortality.

GEORG WOLFF

National Institute of Health

The Theory of Prices, by Arthur W. Marget. New York: Prentice-Hall, Inc. Vol. II. 1942. xxv, 802 pp. $6.00.

To the reviewer's knowledge, this is the biggest and most learned single volume of its kind, surpassing the author's own first volume. Up to 450,000 words are crowded into it, largely in footnotes. Every sentence is strictly "scientific," often of breathtaking length and of painfully precise formulation, with algebraic symbols interspersed, loaded with complex terminologies. Virtually every statement is perfectly logical in itself, and in its very explicit interrelations with others. The whole is supported, in a fashion reminiscent of the Medieval Schoolmen, by about 5,000 imposingly exact and exhaustively analyzed references taken from at least 600 books and articles, in five or six languages, of some 330 authors. Of course, a brief review cannot do justice to such a vast enterprise. Much worth-while detail as well as some controversial judgments must be left without due evaluation.

---

[1] This would come out still more distinctly if deaths from other forms of tuberculosis were also considered; they amount, according to the preceding tabulations of causes of death, to almost one-third (average number in 1929/31: 1,242) of the deaths from pulmonary tuberculosis (average number in 1929/31: 3,879) in Switzerland.

The main outcome of this extraordinary effort is two-fold. First, negative: to prove (by a vast quantity of quotations) the incorrectness of Mr. Keynes' reproach that when economists pass from the general Theory of Value to the Theory of Money and Prices, "we hear no more of these homely and intelligible concepts," namely, of demand and supply; in other words, that Mr. Keynes' claim to be the first one to analyze "output as a whole" in relation to prices is entirely unfounded. Secondly, the positive purpose is to demonstrate (largely by "propositions . . . of an extremely elementary nature") that the Marshallian partial equilibrium analysis is just as useful as it is generally assumed to be. It may be used for monetary theory together with what Marget calls "Stream Equations," i.e., the $MV = PT$ type of quantity-theory formulas, which are called a "translation" of the "Walrasian Process" (p. 384). Briefly, we might go on, undisturbed by Mr. Keynes' claims on originality, operate with individual demand and supply schedules as we did before, and satisfy our sense of formal beauty by knowing that "discrete situations are part and parcel of the process of receiving and expending money in time" (p. 349).

Parturiunt montes: nascitur ridiculus mus. Small as the "mouse" is, it offers endless possibilities for neat formulations and logical exercises. Legions of words are spent to establish the obvious; e.g., that "prices" should mean paid prices, and not imagined ones; that demand curves may be understood "ex-ante," etc. More or less self-evident commonplaces are elaborated by "corollaries" and literary polemics, and substantiated by a plethora of quotations. What might have been fully satisfactory as a short essay—without adding much that is new to our knowledge—is inflated to a large volume, one that does not claim to have any relevance to matters of policy, or to deal with factual analysis. It evades even methodological issues. The "value" of equilibrium analysis, total and partial, is discussed at length but in a dogmatic fashion that misses the essential point: the proper distinction between logical consistency on the one hand, and "realism" on the other.

Mr. Marget's objective is purely "formal." His professed aim is to eliminate the "internecine" struggle between the Professors of Equilibrium Theory and Mr. Keynes by refuting the latter's "unprecedented" attacks against the former.[1] Notice, however, that (a) even though successful, this refutation has no bearing upon any of Keynes' theories of the actual forces underlying monetary phenomena, or upon the policies he pursues, since it deals with the formal categories or "framework" of analysis only, not with analysis itself; (b) Marget's claim that a demand-and-supply approach underlies almost every theory of the value of money since Aristotle, implies a misunderstanding of the problems which Keynes has in mind, and which were so foreign to Aristotle and other "economists" of a pre-capitalistic world as the problems of modern thermodynamics were to pre-19th century physics.

---

[1] Mr. Marget's assumption that no such "internecine" quarrel has ever taken place between (academic) theorists, is altogether erroneous. He apparently overlooks the violent attacks of the Austrian school on the Classics, or of the "Mathematicians" on the "Literary Economists," to say nothing of the fights between Marxist and anti-Marxist theorists.

The reviewer, who labored conscientiously through this heavy volume, cannot refrain from expressing his regrets about the truly uneconomic wasting of Professor Marget's analytical talent and colossal erudition on combating the entirely irrelevant attempts of Mr. Keynes at originality.

Chicago MELCHIOR PALYI

The Theory of Competitive Price, by George J. Stigler. New York: The Macmillan Company. 1942. vii, 197 pp. $3.00.

"This is a preliminary edition of a textbook in advanced economy theory." The scope of the present edition is narrowly circumscribed. Part I—Introduction—includes a discussion of the scope and method of economics, an outline of the functions of the economic system, a discussion of certain basic concepts (competition, statics and dynamics, and equilibrium), and an exposition of the relations between the various quantities used in marginal analysis. Part II—Theory of Competition—covers the theory of demand developed in terms of indifference curves, the nature of costs and the production function, and the pricing of products and productive services under conditions of perfect competition. The author expresses a hope "In the not-too-distant future . . . to add sections on imperfect competition, multiple products, capital theory, and certain other topics which usually receive attention in advanced theory courses."

The author shows a keen sense of balance in the choice between alternative doctrines and a judicious sense of pedagogy in the choice between alternative methods of exposition. The book is simply and concisely written. Mathematical concepts are illustrated by both numerical and geometrical examples. In addition there is considerable symbolic analysis which is conveniently separated from the body of the text. The principal limitation of the book, in addition to the narrow scope of the present edition, is that shared by most texts on modern economic theory, namely that it leaves the impression that there are no open issues. Nevertheless, the reviewer looks forward eagerly to subsequent enlarged editions.

Chapters 7–10, devoted to the nature of costs and the production function and to pricing under competitive conditions, are the best chapters on the subjects available for students. Chapter 9 on pricing under competition relies heavily upon the Marshallian theory of time analysis modified, as has been urged so persuasively by Professor F. H. Knight, by consideration of speculative price. It is unfortunate, however, that greater attention is not given to distinguishing fixed and variable costs. There are many problems attached to this distinction which would be brought to the fore by an operational discussion of costs as seen by the individual firm. A fruitful exposition of short-run normal prices should consider the path by which the individual firm expands and contracts its plant and output. Such a discussion would raise the question of the short-run reversibility of the marginal and average curves and would show the limited usefulness of the concept of the long-run average cost or envelop curve.

Some will question the desirability of approaching the pricing problem via the theory of pure competition, although the author presents a persuasive defense of his choice. Others, including the reviewer, will regret that as now conceived the later editions will evidently ignore many problems of modern theory, including the theory of general equilibrium, the theory of income flow, problems of economic dynamics, and the theory of economic welfare. The theory of demand is an arid subject except as related to the theory of economic welfare or the theory of general equilibrium. The theory of pricing begins to take on significance to most students only when related to the concept of national income or to some so-called "applied" problems. The significance and limitations of partial equilibrium theory are not fully understood until related to the theory of general equilibrium and the theory of income flows under conditions of disequilibrium. Surely such subjects are within the scope of a course in advanced theory. It is hoped that the present conception of the scope of subsequent editions will be extended beyond that now indicated.

J. P. Miller

Yale University

War Economics, edited by Emanuel Stein and Jules Backman. New York: Farrar & Rinehart, Inc. 1942. x, 501 pp. $3.00.

This is a volume which is concerned with the operation of the economic organization in war time and the effects of war upon our economy rather than with the economic causes of war. After noting the stages by which the peace economy is converted into a war economy, the authors proceed to discuss the organization of industrial production, agriculture, transportation, electric power, and foreign trade for the conduct of the war. Other chapters deal with the problems of war financing, price control, housing, business finance, and consumption. The problems of labor also receive attention. The closing chapter discusses the problems of industrial demobilization after the war.

The volume provides a clear and convenient outline of the economic problems attending total war and studies the various alternative approaches to these problems. The necessity of planning for total war is emphasized, but the problems which will follow in the wake of the global conflict are not neglected. Much useful information has been assembled by the authors which is descriptive of the steps taken in this and other countries in the effort to prosecute the war to a successful conclusion. Fortunately for us, many lessons may be learned from the experience of the other belligerents.

As one reads the volume he gains an appreciation of the immensity and complexity of the battles of production, transportation, financing, and marketing. It should prove helpful and interesting to students of economics and should provide a ready text and reference for those who wish to inform themselves of measures taken in organizing our economy for the war and the problems which face us. A wide range of topics is covered in a book which is not large but which contains a great deal of information.

On particular points the reader may wish to disagree or he may feel that some favorite theory has been neglected. Not every topic of importance in the economics of war could be treated, but the authors have succeeded in presenting a surprisingly comprehensive description and discussion of our economy in war time.

ALFRED G. BUEHLER

University of Pennsylvania

The Women Graduates of a Collegiate School of Business, by Ann Brewington and Evelyn Van Emden Berg. Chicago: The University of Chicago Press. Studies in Business Administration, The School of Business. 1942. ix, 99 pp. $1.00.

"The purpose of this study is to record some of the occupational and personal experiences of the women graduates of a collegiate school of business." A twelve page questionnaire was mailed to the women graduates of the University of Chicago's School of Business. Herein are recorded the replies of 245 of the 462 women who were graduated between January 1913 and January 1941. The recording takes some sixty pages of text, thirty-five pages of appendix, and patience.

The reporting group seems to be representative of the known characteristics of the original 462 graduates at the time of graduation, no claim being offered that the group is representative of anything more than these graduates. The result is a listing of miscellaneous information, of interest primarily to those persons who are curious about the women graduates of the University of Chicago's School of Business. Even here, however, generalization is difficult, for the findings are often uncontrolled for such primary refinements as age at graduation, and date of graduation.

Five chapters are presented under the rubrics "Educational Preparation of the Women Studied," "Careers," "Finances," "The Personal Equation," and "A Long Look Ahead." Each chapter intones a procession of absolute and relative figures calculated to impress the reader with the unexceptionable character of these girl graduates, their varied careers, sundry financial conditions, variegated personalities, and the fact that they face the future with mixed feelings of hope and fear.

The reader may judge from these glyptic globules just how well the obvious has been elaborated:

> "The 53 homemakers have been married an average of 7.9 years, and their average age is 34 years. . . . The value of their School of Business training is evidenced by the fact that some occasionally assist their husbands . . . 24 are always consulted by their husband about business problems, and 18 are sometimes consulted. . . ."
>
> "College education in business widens employment opportunities."
>
> "Application of professional knowledge and skills to insignificant and boring routines in business, as well as in housekeeping, results in happiness."
>
> "Salaries increase with experience."
>
> "Women should help one another by maintaining a cooperative attitude toward one another."

It would be generous to close with this last bit of wisdom: "This is not a man-made world but a mismade world, and the fault is not entirely that of man."

DAVID YENTIS

Federal Public Housing Authority

How Collective Bargaining Works. New York: The Twentieth Century Fund. 1942. xxviii, 986 pp. $4.00.

According to the foreword, "This volume contains full-length sketches of the actual workings of collective bargaining in sixteen United States trades and industries and thumbnail summaries covering thirteen other fields." Sixteen authors have cooperated, under the direction of Dr. H. A. Millis, to provide fourteen industry and trade essays. Philip Taft has, in addition, summarized the development of collective bargaining in this country and outlined its present status in thirteen additional industries. A final section of the appendix presents available data as to membership of A. F. of L., C.I.O., and unaffiliated national and international unions.

A fairly consistent pattern is discernible in the treatment accorded each of the industries. Authors first present a more or less comprehensive description of the economic characteristics of each industry, in terms of volume of output and employment, location, competitive features, special technological conditions, major labor problems, and any distinctive economic aspects or notable recent trends. There follows a description of the development of collective bargaining in the industry, with special attention to historic struggles, the structure, functions, and present strength of unions involved, and to general and distinctive features of trade agreements developed or in force in the industry. Employers' associations, if any, receive attention at this point. The next section considers the results of collective bargaining, with special reference to changes in working hours, rates of pay, the settlement of disputes, special work rules, effects of collective bargaining on unit costs, and union cooperation in meeting competition. A final division is devoted to appraisal, by the author or authors, of the over-all effects of collective bargaining in the industry, including a description of unsolved problems.

The studies thus presented will provide valuable background and perspective for all students of modern industrial relations. They are not uniformly well done, and no single criticism applies equally to all of them. Several of the industry studies tend, in the opinion of this reviewer, to forget that unionists remain a distinct minority in the nation's labor force, and that collusion with employers may advance union members' interests at the direct expense of the public. One lesson appears clearly; that is the utter necessity for intelligent, broadminded, powerful union leadership if collective bargaining is to become anything more than the socially dangerous tool or device of a selfish interest group.

DALE YODER

University of Minnesota

Determining Work Loads for Professional Staff in a Public Welfare Agency, by Herbert A. Simon et al. Berkeley: University of California, Bureau of Public Administration. 1941, xii, 94 pp. $1.00.

This monograph describes a study to determine the size of workload in a public assistance agency which will result in the highest proportion of eligible cases, together with satisfactory service to recipients, at the lowest total cost. The study was conducted in two districts of the Los Angeles County Emergency Relief Administration by the Bureau of Public Administration of the University of California.

The experiment studied three steps in the performance of the agency function:

1. The intake or "qualifier" interview.
2. Home and collateral visits by the field intake worker.
3. The reinvestigation of active cases or "carrier" load.

Careful controls were maintained to establish comparable conditions in the workloads which varied in size from the current average in the agency to points substantially above and below it. Of special interest was the adjustment period of a month allowed the investigators with carrier loads so that no backlog of work would distort the findings. The maintenance of the carrier loads at a fixed number offered a stability in workload which is seldom duplicated in actual public welfare practice and may therefore be questioned.

The most reliable test of eligibility was considered to be in the number of cases rejected as ineligible after acceptance. Results were measured by the analysis of special statistical forms kept on each case, by a social audit on a sample basis which included analysis of case records and visits to recipients and collaterals and by a time study. To this reviewer, the visits to clients and collaterals by the social audit group as a check on the experiment seem both costly and socially undesirable.

The findings were more satisfactory in measuring the determination of eligibility than the adequacy of service. The latter was assessed largely in terms of referrals to other agencies. Whether lower workloads resulted in better worker-client relationship, and therefore, a more constructive experience for the client, is difficult to ascertain and this was recognized by the study group.

An interesting aspect of the findings is that the intake or qualifying interview showed the greatest gain with a lower workload, the active, or carrier load, was next, and the intake visiting in the field showed the least gain with lower loads. This again emphasizes the value of careful intake interviews and the average of 30 to 40 minutes allowed with the low workload is not impractical.

The findings of this study agree with those of similar studies in other parts of the country, namely, that lower workloads than usually exist in public welfare agencies would cost less and offer better service to agency recipients. Most public welfare administrators know this from practical experience without the benefit of a scientific study but factors beyond their control often prevent their acting on this knowledge. The size of workload for maximum efficiency in most public assistance agencies can perhaps be determined as satisfactorily on the basis of experience, observation and analysis of local factors affecting size of workload as by the more elaborate and costly study method.

The Los Angeles study has a particular contribution in this field because it so carefully details the methods used and the various elements which must be considered in reaching the findings. This emphasis on methodology should render it particularly valuable to the public welfare administrator who finds it necessary to conduct a similar study.

Constance Hastings

Social Security Board

Random Sampling Distributions, by Alan E. Treloar. Minneapolis: Burgess Publishing Company. 1942. 94 pp. $2.25.

This book is a planographed set of notes rather than a formal textbook. It is intended as an aid to the serious student of science who, requiring significance test techniques, is neither acquainted with their derivation nor satisfied to use formulas blindly. With this reader in mind, the author relies upon simple algebraic formulations of the argument, largely in terms of the moments of the distributions, and a liberal appeal to the geometry of the problems, presented in 24 drawings. Since one or both of these approaches is within the grasp of most science students, the book should prove to be a real contribution to them.

The distributions dealt with are those concerned with significance tests on means, standard deviations, and correlation coefficients. In addition to the text discussion around the development of the relevant distributions, the four following appendix tables are given: a table for estimating σ from s on the basis of assumptions concerning the position of s in its random sampling distribution; 5 per cent and 1 per cent points for the distribution of F (reproduced from Snedecor's text); table of the probability that t, based on n degrees of freedom, will be exceeded; tables of z_r as a function of r (reproduced from the author's text).

Of particular value to the applied student approaching these significance tests is the emphasis given to the assumptions from which the algebra flows. These are not only stated explicitly but the limitations which they place upon the interpretation are stressed. It is pointed out that most of our exact sampling distributions rest on an assumption that the universe sampled is normal (an assumption that cannot be verified in most actual problems) and therefore the sampling distributions may be only approximate for the problem in which they are used. Thus the author questions "the value of parading that exactness feature of the test when it rests on an assumption that is an approximation."

Various controversial issues with regard to the importance of the significance test in statistical reasoning, and the interpretation of the probability attached to a given test are discussed. With regard to the latter, Dr. Treloar does not allow a fractional probability to be attached to the statement that two parameters differ, or that a given parameter lies within a specified range. This probability he allows to be either 0 or 1. Rather, he uses the probability in a subjective sense, to measure the degree of confidence we have in the truth of the statement. However, in the discussion of the null hypothesis, this interpretation does not seem to be strictly maintained, for we have the statement, "If that probability is less than 50 per cent then the odds may be said to favor the difference *not* having arisen from random sampling, and one is therefore disposed to cast doubt on the accuracy of the null hypothesis." The first part of this statement is an inverse probability statement, yet without any consideration of existence probabilities. It seems to be out of line with the rest of the discussion of interpretation of significance tests.

The student will find this a book which not only lays a substantial groundwork for the tests he has frequent occasion to use, but also one in which real consideration rather than a pat answer is given to the meaning of the tests.

MARGARET MERRELL

School of Hygiene and Public Health
The Johns Hopkins University

Poisson's Exponential Binomial Limit, by E. C. Molina. New York: D. Van Nostrand Company, Inc. 1942. 47 pp. $2.75.

The book contains three pages of introduction, one of which, explaining some inspection problems, is signed by H. F. Dodge; and 47 pages of tables. These are divided into two parts. Table I gives values of the individual terms of the Poisson Law,

$$p(x\,|\,a) = e^{-a}a^x/x!$$

and Table II the values of the probabilities of a Poisson variable being at least equal to c; thus

$$P(c,\ a) = \sum_{x=c}^{\infty} p(x\,|\,a).$$

Beginning with $a = .001$ and up to $a = .010$ the tables proceed by increments of a equal to .001. Further on and up to $a = .30$ the increment is .01. Still further and up to $a = 15.0$ the increment is .1 and beyond $a = 15.0$ it is one unit. The last columns of both tables correspond to $a = 100$. The number of decimals given is also varying. Up to $a = .40$ the figures given are with seven decimals, beyond that limit with six.

As far as the reviewer is aware the present table is more extensive than any other table of the Poisson Law in existence. In spite of the wider and wider applications of the Poisson Law, its available tables were limited to those given in K. Pearson's *Tables for Statisticians and Biometricians*, Part I. Of the tables given there, Table LI gives the values of individual terms of the Poisson Law for $a \leq 15$, proceeding by increments of .1 and Table LII the sums of such terms, for $a \leq 30$, proceeding by increments of 1. However, the arrangement of Table LII, meant for one specific purpose, is somewhat awkward to use in other cases.

Of course the numerical values of the Poisson Law for higher values of a can also be obtained from the extensive *Tables of the Incomplete Γ-Function* edited by Karl Pearson (Biometrika Office, London, 1934). However, that table was again mainly meant to serve purposes other than applications of the Poisson Law, namely those encountered in work with the Pearson's Type III frequency curve. These applications and also difficulties in setting up the table have forced Pearson to use a system of arguments which contributes somewhat to the inconvenience in using the *Tables of the Incomplete Γ-Function* for purposes of computing the values of the Poisson Law.

As a result of all these circumstances, the volume under review is a most welcome addition to the growing arsenal of statistical tools. There is no doubt whatsoever that in a relatively short time it will find its place on the shelves and desks of all statistical workers, exhausting the first edition. We may hope that then the Van Nostrand Company will become convinced of the importance of this particular publication and will re-issue it in a proper printed form. The present edition is lithoprinted and, in places, regretfully, none too well.

J. NEYMAN

University of California

Encyclopaedia Britannica World Atlas. Chicago: Encyclopaedia Britannica. 1942. Vol. XI. 223 map plates, 272 pages of text and tables, 128 pages of map index.

Although the publication of this atlas by the Encyclopaedia Britannica marks an event of interest in the field of cartography the volume is primarily noteworthy for the scope and quality of its statistical contribution, a contribution representing a new departure for this estimable research establishment. For this innovation credit is due G. Donald Hudson of Northwestern University. This timely document merits careful study by those who may exert, or hope to exert, influence in rehabilitating the war-torn world. Only by giving conscientious consideration to such carefully compiled world-wide factors can a lasting peace be built.

This atlas provides a useful series of world, hemisphere, regional, national, and, for the United States, state-county political maps showing place names, boundaries, lakes, rivers, canals, and railroads. In addition there are eight eye-pleasing physical maps showing elevations by tint bands. These political and physical maps, comprising over 200 plates with a number of large scale inserts, were prepared for Britannica by C. S. Hammond & Company to show the status of the world on September 1, 1939.

"Geographical summaries" gives the new statistical setting. The year 1938 was selected for the base, although statistics are provided when possible for subsequent years. This tabular presentation, covering 237 atlas sized pages, divides the world into seven continental regions, namely, North, Middle and

South America, Europe, Asia, Africa and Australasia and the Pacific Islands. For each region a short page or two of descriptive orientation introduces the statistics which follow, statistics carefully selected through an exhaustive examination of governmental and institutional sources. Over 500 of these sources are detailed in valuable regional bibliographies. The first series of nine tables group all countries of a region together. Only the very smallest countries are omitted. These tables provide ready comparisons between countries. The following information is given by countries grouped for each region: area, population and type of government; leading agricultural, mineral and manufactured products; dollar value, major commodity and major source of imports and exports; communications facilities; transportation facilities; value of currency in dollars; government expenditure for military purposes; military service and personnel; and military equipment and organization. The regional tables are followed by fifteen more for each of 165 countries, or approximately 2,400 individual tabulations.

The atlas is concluded with a short section of 22 pages containing maps and compilations of data for "world spheres of influence," or what might have been more appropriately called "areas of control." Spheres of influence are shown individually as they existed for the United States, Great Britain, France, Germany, Spain, Japan, Italy and Portugal on August 1, 1942. The usefulness of this series of maps and accompanying data might have been greatly enhanced if they had conformed to a more significant date. Though it might be desirable to make all sections of an atlas as current as possible, still it seems unlikely that August 1, 1942, the last possible date prior to publication, will have any long run significance. Compilations whether cartographic, statistical or explanatory when keyed to a date falling into the uncertain turbulence of the rapidly changing events of a world war, partake a momentary character akin to a newspaper statement. This unfortunately is true no matter how painstakingly accurate the treatment of the material. It would seem, therefore, that the choice of a date just prior to the partitioning of Czechoslovakia or the invasion of Poland by Germany might have resulted in a series of maps and compilations destined to serve more significantly as a reference base in planning and intelligently appraising the peace settlement at the close of this war.

M. J. PROUDFOOT

War Production Board

Distribution of Ownership in the 200 Largest Nonfinancial Corporations, by Raymond W. Goldsmith, Rexford C. Parmelee, Irwin Friend, James Gorham and Helene Granby. Temporary National Economic Committee. Monograph No. 29. Washington. xviii, 1,557 pp.

Survey of Shareholdings in 1,710 Corporations with Securities Listed on a National Securities Exchange, by Helene Granby. Temporary National Economic Committee. Monograph No. 30. Washington. xviii, 258 pp.

Monograph 29 aims to show the degree of concentration of stock ownership in the largest American corporations. Only 144 of the 1,557 pages are devoted to a discussion of the material, though some additional comments appear in Appendix I and II. The rest of the book consists of tables showing the size of stockholdings in the 200 largest corporations. In all, 408 security issues are analyzed. Much of this information is in the form of frequency distributions, showing the number of shares held in each of many different classes representing the size of the stockholding. These distributions are prepared both with classes representing the numbers of shares per shareholding and with classes representing their value. Succeeding tables analyze these data from many points of view.

There are also tables showing the number of shareholders that are individuals, corporations, trusts, etc., while another section shows the details of the shareholdings of officers and directors of each of the 200 corporations. An interesting set of tables shows the names of the twenty largest shareholders of each of these corporations; this is shown both for the twenty largest shareholders of record and for the twenty largest legal and beneficial shareholders. Data are presented on shareholdings of the Mellon, Rockefeller, Du Pont, and other families.

The text is devoted to discussion of the sources of data, methods of analysis, and conclusions and "implications." One of the most interesting parts of the text is the comparison of the degree of concentration of ownership of different companies in the same industrial division. For each of many branches of industry, graphs are drawn presenting Lorenz curves for individual corporations.

The degree of concentration of ownership as shown by these curves and other data appearing throughout the text is indeed pronounced. It is pointed out that it is not the directors or management that hold this high degree of power, but usually the family of the founders—in its own name and in trusts, estates, and personal holding companies. In the case of most families, the fortune has been retained through reinvestment in the securities of the original company. Branching out into other lines of enterprise is noted only in the case of the Mellon family.

The monograph is well worked out as a study of the concentration of power. It is also valuable for students of corporation finance and investments. There is a little too much emphasis on statistics as such; as with other TNEC studies, the data themselves do not always point to as many horrors as we suspect the authors would have us believe.

Monograph 30 "provides a background for, rather than a supplement to, the companion study . . ." It is an analysis of the 1,710 corporations, financial and nonfinancial, which have securities listed on national security exchanges, whereas the preceding monograph studied the 200 largest nonfinancial corporations whether or not listed. The analysis of concentration of shareholdings is comparable with that in Monograph 29. However, because of the larger number of cases, it was possible in many instances further to subdivide the industries into which the companies fall. Also, subdivisions were made to disclose the distribution of shares in widely-held corporations as contrasted with closely-held corporations.

As in the previous monograph, the tabular presentation takes up by far the larger portion of the book. Lorenz curves corresponding to those in Monograph 29 are used to portray the concentration of stockholdings. Instead of a curve for each corporation, however, only curves representing groups of corporations are presented. These groupings, corresponding to classifications in the tabular material in the appendix, are based on the industry with which the corporation is identified and on its size as measured by amount of assets.

Throughout this monograph the emphasis is on totals rather than on individual companies. Logically, Monograph 30 should have come first; as the introduction says, the study provides a broad basis on which analysis of individual companies, their directors, and their shareholders is built in Monograph 29.

RALPH C. EPSTEIN
JOHN H. MYERS

The University of Buffalo

Merchandise Availability in Utah, by David E. Faville, Dix M. Jones and Richard B. Sonne. Stanford University: Stanford University Press. 1942. xvi, 229 pp. $3.00.

This is a study of merchandising in rural Utah. It is the result of a survey conducted by the Stanford Graduate School of Business at the request of a "far-seeing group of Western Merchants (and with their financial backing) to ascertain whether similar income groups living in different localities, especially in small villages as contrasted with larger towns, are enabled to enjoy to the same degree the various comforts of life, or whether in certain districts higher prices and a limited range of merchandise result in various scales of living." No particular area was designated as the field for study, but Utah was selected primarily because of its distinctive rural villages, its population characteristics, and its western location.

The report deals mainly with the availability and retail price variations of commonly used food, drugs, clothing, dry goods and miscellaneous items. It tells where and why different groups did their purchasing, whether purchases were made locally or out of town, and for what reasons; whether the stores were independent, corporate chains or contract chains. The authors tell us whether or not the stores surveyed were neat, clerks courteous, service prompt; whether there was a good selection of merchants available and whether prices for advertised goods were uniform in competitive stores in the same community.

The original study was conducted by the Survey Method in which housewives and store operators in selected rural communities were interviewed. In all, 2,156 consumer interviews were obtained over a period of four months, 1,122 in the rural towns and 1,034 in the control cities. This represented approximately a 25 per cent sample of both areas. The survey also included interviews with 170 store operators in the four control cities and 57 in rural towns plus detailed checking of shelf prices at the stores. The questionnaires thus obtained were then tabulated on International Business Machine cards. These cards were sorted according to various factors and the results analyzed statistically.

The report, therefore, contains many interesting tables and charts designed to reveal the vast amount of significant findings that came out of the survey.

In the price section some attempt was made to determine the reliability of price variations by correlation and significant differences technique. Some interesting relationships were obtained where sufficient data were available. However, in some cases apparent differences proved to be insignificant; hence no reliable conclusions could be drawn.

The more important conclusions were as follows: "In so far as the ready availability of a wide range of drug, variety, clothing, and dry-goods articles bears on the enjoyment of the various comforts of life, the residents of small rural towns and villages interviewed in this survey were at disadvantage as compared with their neighbors in medium sized towns. Limited selection, high prices, or complete unavailability were common for these items in the rural districts, and these deficiencies were most keenly felt by farmer and laborer groups." They also concluded that "the rural-town stores surveyed did not commonly carry a variety of fresh fruits, vegetables, and meats."

The survey appears to have been well planned and ably executed. The results were thoroughly analyzed and written up in an interesting style. The few criticisms heard by the writer from others who have read the book might be summed up in the slang phrase, "so what?" In other words, the job was well done and the story well told, but relatively little new light has been thrown on the screen.

DILWORTH WALKER

University of Utah

PUBLICATIONS RECEIVED*

Bard, Erwin Wilkie. *The Port of New York Authority*. New York: Columbia University Press. 1942. x, 352 pp. $3.50.

Barger, Harold. *Outlay and Income in the United States, 1921–1938*. New York: National Bureau of Economic Research. 1942. xxvii, 391 pp. $2.50.

Beveridge, Sir William. *Social Insurance and Allied Services*. New York: The Macmillan Company. 1942. 299 pp. $1.00.

Brady, Robert A. *Business as a System of Power*. New York: Columbia University Press. 1943. xx, 340 pp. $3.00.

Branch, Melville C. *Urban Planning and Public Opinion. National Survey Research Investigation*. Princeton: The Bureau of Urban Research, Princeton University, Research Series No. 1. 1942. 87 pp. $1.00.

Chase, Stuart. *Goals for America*. New York: The Twentieth Century Fund. 1942. viii, 134 pp. $1.00.

Commercial Policy in the Interwar Period: International Proposals and National Policies. Trenton: Trenton Printing Company, League of Nations. 1942. 164 pp. $1.75.

Commodity Statistics, The 1942 Commodity Year Book. New York: Commodity Research Bureau, Inc. 1942. 294 pp. $5.00.

Commodity Year Book. New York: Commodity Research Bureau, Inc. 1942. 413 pp. $7.50.

Davis, Hiram S. *Wool and the War*. Philadelphia: University of Pennsylvania, Wharton School of Finance and Commerce, Research Report No. 6. 1942. xi, 107 pp. 50 cents.

Fabricant, Solomon. *Productivity of Labor in Peace and War*. New York: National Bureau of Economic Research, Occasional Paper 7. 1942. 28 pp.

Farnsworth, Helen C. *Wartime Food Developments in Germany*. Stanford University, California: Food Research Institute. No. 1. 1942. 36 pp. 25 cents.

Guide for the Study of American Social Problems. New York: Columbia University Press. vii, 181 pp. $1.00.

Guilford, J. P. *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill Book Company, Inc. 1942. xi, 333 pp. $3.25.

Hall, Marguerite F. *Public Health Statistics*. New York: Paul B. Hoeber, Inc. 1942. xxii, 408 pp. $5.50.

Heindel, Richard H. *The Integration of Federal and Non-Federal Research as a War Problem*. Technical Paper Number 9. Washington: National Resources Planning Board. 1942. iv, 122 pp.

Hill, A. Bradford. *Principles of Medical Statistics*. London: The Lancet Limited. 1942. vii, 189 pp. 7/6d net.

Katona, George. *War without Inflation*. New York: Columbia University Press. 1942. x, 213 pp. $2.50.

Kiser, Clyde V. *Group Differences in Urban Fertility*. Baltimore: The Williams & Wilkins Company. 1942. xii, 284 pp. $2.50.

Kuczynski, Jurgen. *A Short History of Labour Conditions in Great Britain and the Empire*. London: Frederick Muller, Ltd. 272 pp. 12/6d.

Millspaugh, Arthur C. *Peace Plans and American Choices*. Washington: The Brookings Institution. 1942. vii, 107 pp. $1.00.

The Network of World Trade. Princeton: Princeton University Press. League of Nations Publication. 1942. II. A. 3. 172 pp. $2.75.

Nourse, Edwin G. *Distribution Costs—Waiting at the Bunghole*. Washington: The Brookings Institution. 1942. v, 25 pp. 25 cents.

---

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. Some items of minor interest to statisticians have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications which are to be found in the *Revue de l'Institut International de Statistique, Journal of the Royal Statistical Society, American Economic Review, Population Index, Transactions of the Actuarial Society of America, The Record of the American Institute of Actuaries*, and *Sankhyā, The Indian Journal of Statistics*.—Editor.

Nourse, Edwin G. *Wages as Cost and as Market.* Washington: The Brookings Institution. Pamphlet No. 44. 1942. ix, 43 pp. 25 cents.

Palsgrove, Grant K. *Solution of Special Problems in Pipe Flow by Graphical Analysis.* Troy: Rensselaer Polytechnic Institute Bulletin, No. 58. 1942. 31 pp.

Payne, Wilson F. *Business Behavior 1919–1922. An Account of Post-war Inflation and Depression.* Chicago: The University of Chicago Press. xi, 215 pp. $1.50.

Pearson, James Constantine. *The Reciprocal Trade Agreements Program: The Policy of the United States and Its Effectiveness.* Washington: Catholic University of America Press. 1942. xiii, 328 pp. $2.00.

Peters, J. *Seven Place Values of Trigonometric Functions for Every Thousandth of a Degree.* New York: D. Van Nostrand Company, Inc. 1942. $7.50.

The Political Economy of Total War. Washington: School of Foreign Service, Georgetown University. 1942. 128 pp.

Refugee Settlement in the Dominican Republic. A Survey Conducted under the Auspices of the Brookings Institution. Washington: The Brookings Institution. 1942. xvi, 410 pp.

Report to the Council on the Work of the Joint Session. Princeton: Princeton University Press. League of Nations Publication. 1942. II. A. 4. 25 cents.

Scott, James Dacon. *Advertising Programs for Products with Selected Distribution.* Boston: Harvard University Graduate School of Business Administration, Bureau of Business Research. 1942. v, 112 pp.

Scoville, John, and Sargent, Noel. *Fact and Fancy in the T.N.E.C. Monographs.* New York: National Association of Manufacturers of the United States of America. 1942. xvi, 812 pp.

Tax Yields: 1941. Philadelphia: Tax Institute, University of Pennsylvania. 1942. xi, 171 pp. $2.50.

von Mering, Otto. *The Shifting and Incidence of Taxation.* Philadelphia: The Blakiston Company. 1942. xiii, 262 pp. $3.50.

Wartime Accounting, 1942. Papers Presented at the Fifty-Fifth Annual Meeting of the American Institute of Accountants. New York: American Institute of Accountants. 1942. v, 182 pp. $1.00.

Wartime Rationing and Consumption. Economic Intelligence Service. New York: Columbia University Press, League of Nations. 1942. 87 pp. $1.00.

Whitaker, Arthur P., ed. *Inter-American Affairs, 1941.* New York: Columbia University Press. 1942. vii, 240 pp. $3.00.

Woodbury, Robert Morse. *Food Consumption and Dietary Surveys in the Americas.* Montreal: International Labour Office. 1942. iv, 64 pp. 35 cents.

Government Publications

Area Analysis. A Method of Public Works Planning. Technical Paper Number 6. Washington: National Resources Planning Board, U. S. Government Printing Office. 1942. iv, 31 pp. 15 cents.

Arquivos de Higiene. Rio de Janeiro: Imprensa Nacional, Ministerio da Educação e Saude, Departament Nacional de Saude. 1942. Año 12, 300 pp.

Atividades Estatisticas. Santa Catarina: Instituto Brasileiro de Geografia e Estatistica. 1942. 55 pp.

Boiteux, Henrique. *Madeiras de Construção de Santa Catarina.* Publicação N. 27. Santa Catarina: Imprensa Oficial do Estado. 1942. 107 pp.

Cabral, Oswaldo R. *Medicina, Medicos e Charlataes do Passado.* Florianopolis: Instituto Brasileiro de Geografia e Estatistica, Departamento Estadual de Estatistica. 1942. Publicação N. 25. 292 pp.

Censo Nacional de 1940, Resultados Generales. Lima, Peru: Ministerio de Hacienda. 1941. 68 pp.

Distribution of Manufacturers' Sales, 1939. Washington: United States Government Printing Office. Census of Business, Volume 5. xii, 206 pp. $1.25.

Employment Security in Ohio, 1941. Columbus: State of Ohio Bureau of Unemployment Compensation. 1942. 81 pp.

Estado de la Instrucción en el Perú según el Censo Nacional de 1940. Lima, Peru: Imprenta Torres Aguirre. 1942. 67 pp.

Extracto Estadistico Del Perú. Lima, Peru: Imprenta Americana, Ministerio de Hacienda y Comercio. 1942. 748 pp.

Griffin, Grace Gardner, Louraine, Dorothy M., Patterson, Margaret K. *Writings on American History, 1937 and 1938.* 1942. Washington: Government Printing Office. xxxvi, 869 pp. $1.75.

Indicador Agro-Pecuario Industrial, Comercial, Bancario e Postal-Telegrafico. Florianopolis: Instituto Brasileiro de Geografia e Estatistica, Departamento Estadual de Estatistica. 1942. Ano I, 114 pp.

National Resources Planning Board. *The Future of Transportation.* September, 1942. 43 pp.

The Paper and Pulp Industries of New York State. Albany: Executive Department of State of New York, Publication No. 2. 1942. 111 pp.

Peluso, Vitor A., Jr. *Rio Do Sul.* Santa Catarina: Instituto Brasileiro de Geografia e Estatistica, Departamento Estadual de Estatistica. Publicação N. 26. 133 pp.

Recopilacion Estadistica, Años 1939–1940. Buenos Aires: Talleres Graficos de la Penitenciaria Nacional. 1942. 711 pp.

Review of the Trade of India: in 1940–41. Calcutta: Printed by the Manager, Government of India Press, Sixty-seventh Issue. 1942. 287 pp.

Tax Delinquency and Rural Land-Use Adjustment. Technical Paper Number 8. Washington: National Resources Planning Board, U. S. Government Printing Office. 1942. iv, 190 pp. 50 cents.

Welch, Charles H., and Sevin, Charles H. *Record Keeping for Small Stores.* Washington: U. S. Government Printing Office, Special Committee to Study Problems of American Small Business. 1942. 93 pp.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

June, 1943 · $1.50 per Copy · $6.00 per Annum · Vol. 38 · No. 222

CONTENTS

ARTICLES

Factors in the Variability of Response in Enumerative Studies . . . . Gladys L. Palmer 143

Time Series Significance Tests Based on Signs of Differences . . . . Geoffrey H. Moore and W. Allen Wallis 153

An Approach to the Measurement of Farm Population Pressure in Wisconsin . . . . Thomas C. McCormick 165

The Production Function for Canadian Manufactures . . . . Patricia Daly and Paul H. Douglas 178

The Composition of Income and Ownership of Capital by Income Classes in the United States in 1936 . . . . Rufus S. Tucker 187

Comparison of Demographic Rates Assumed by the National Resources Committee with Actual Experience . . . . Robert J. Myers 201

Choice of the Dependent Variable in Regression Analysis . . . . Frederick V. Waugh 210

Comments . . . . Mordecai Ezekiel 214

Note on a Common Statistical Inequality . . . Irving H. Siegel 217

Contents of Tippett's "Random Sampling Numbers" . Robert Gage 223

Quality Control Applied to Business Administration . . Wm. B. Rice 228

Family Versus Individual in Measurement of Audiences Alfred Politz 233

The Cost of Living in Cities in Relation to Their Size and Latitude . . . . E. L. Thorndike and Ella Woodyard 238

Experience with Tests of Significance: A Reply to Professor R. A. Fisher . . . . Joseph Berkson 242

Edward Lewis Dodd, 1875–1943 . . . . C. D. Simmons 247

The Study of Cycles 248

A Correction 249

BOOK REVIEWS

Beebe, Gilbert Wheeler. *Contraception and Fertility in the Southern Appalachians.* Margaret Jarman Hagood 252

Bennett, Kenneth R., see Pearson, Frank A.

Brookings Institution. *Refugee Settlement in the Dominican Republic.* Erich W. Zimmermann 254

Commodity Statistics, The 1942 Commodity Year Book. Oscar L. Endler . . 260

Commodity Year Book. Oscar L. Endler 260

Culliton, James W. *Make or Buy.* Charles E. Young 261

Davis, Hiram S. *Wool and the War.* R. R. Renne 262

Guilford, J. P. *Fundamental Statistics in Psychology and Education.* Martin A. Brumbaugh 266

Haberler, Gottfried. *Consumer Instalment Credit and Economic Fluctuations.* James F. Cusick 256

Continued on next page

Hall, Marguerite F. *Public Health Statistics.* Clyde V. Kiser 253
Holmes, Maurice C. *An Outline of Probability and Its Uses.* E. C. Molina 267
Horton, Donald C., Larsen, Harald C., and Wall, Norman J. *Farm-Mortgage Credit Facilities in the United States.* John B. Schneider . . 257
Kiser, Clyde V. *Group Differences in Urban Fertility.* Frank Lorimer . . 250
Larsen, Harald C., see Horton, Donald C.
League of Nations. *Wartime Rationing and Consumption.* M. K. Bennett 264
Payne, Wilson F. *Business Behavior 1919–1922.* V. P. Timoshenko . . 258
Pearson, Frank A., and Bennett, Kenneth R. *Statistical Methods Applied to Agricultural Economics.* E. L. Butz 265
Peters, J., compiled by. *Seven-Place Values of Trigonometric Functions.* Kenneth N. Ogle 268
Wall, Norman J., see Horton, Donald C.

Publications Received 269

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 222 JUNE, 1943 *Volume 38*

FACTORS IN THE VARIABILITY OF RESPONSE IN ENUMERATIVE STUDIES

BY GLADYS L. PALMER
Industrial Research Department, University of Pennsylvania

THERE ARE SEVERAL sources of error in enumerative studies. The error that arises in sampling is well recognized, but other sources, which may be more important, are often overlooked, even though they are present as much in complete coverages as in samples. For instance, there is a variability of response to be reckoned with even within the framework of well defined schedule concepts and well trained enumerators. Poorly worded schedule questions and inadequate training of enumerators not only intensify this variability, but are the cause of unreliability in the sense of biases. However, in the absence of a standard operation, which by definition would yield no error and would enable one to differentiate correct from incorrect results, and thus to recognize biases, it is perhaps better to speak of variability of response merely as a source of difference between the results of two or more coverages of the same universe, rather than to speak of a bias in one or both coverages. The purpose of this paper is to present some results on variability of response.[1]

Relatively little is known about the degree of variability of response for different types of information or its relationship to other conditions of enumeration because the opportunity to measure it is limited to studies in which two sets of interviews requesting the same information for the same persons for the same period of time have been made. Relatively high costs have discouraged this kind of enumerative study while unfortunately, at the same time, differences in the results from applications of parallel techniques to the same universe for measuring the same phenomenon, without adequate explanation of the causes of these differences, have led to public confusion. As a consequence, doubt has been cast on the results of enumerative studies, yet these are the only, or the major, source of information for many types of important social and economic data.[2]

---

[1] The writer is indebted to Dr. W. Edwards Deming and Dr. A. Ross Eckler of the Bureau of the Census for helpful suggestions with regard to this analysis.

An opportunity to study variability of response is found in two sets of interviews of 8519 persons in 2686 Philadelphia households obtained by two groups of enumerators who had been subject to the same training program. The purpose of this analysis is to determine the relation of certain conditions of enumeration to variation in individual response and the effect of variation in individual response upon the summarized results.

Four important social and economic characteristics were selected for the analysis of variability of response: marital status, age, employment status, and education. These characteristics were related to the following conditions of enumeration:

1. The proportion of Negro or foreign-born residents, or both combined.
2. The number of days of experience of enumerators at the time of the first field visit.
3. The lapse of time between the two interviews.
4. The personnel ratings of the original and revisit enumerators.[3]

Findings of this Study.—The findings of this study indicate that in the census type of survey, where the housewife is the usual respondent, one may expect a negligible variation response for marital status (1.7 per cent), a higher variation in response for employment status (9.9 per cent), and age (12.8 per cent), and the highest for education (23.4 per cent). These differences in variability of response arise partly from inherent difficulties in identification of status or classification. This is especially true with respect to returns on employment status for persons whose labor market attachments are difficult to identify or classify. It is also true for information on education for persons whose schooling has been informal or is described in terms of the educational systems of foreign countries which differ from the American school system. Variation in response for age or marital status cannot be explained in these terms, however, and is apparently the result of indifferent or inaccurate reporting.

Variability of response is lower for returns from districts in which there is a relatively high proportion of native-white residents than for districts in which there is a relatively high proportion of Negroes or of Negroes and foreign-born white residents combined.

---

[2] For a summary of recent studies discussing this problem, see Gladys L. Palmer, *The Reliability of Response in Labor Market Inquiries*, Technical Paper #22, *Report to Federal Statistical Agencies*, Bureau of the Budget, July 1942.

[3] The personnel rating was based on all work of the enumerators during the survey and not on the results of the revisit program. Ratings were made by supervisors for personal appearance and personality, accuracy, learning ability, and output. These were averaged to give grades A, B, C, D, E, and F and yield a rough estimate of the comparative ranking of the enumerators' work.

None of the other conditions of enumeration which were selected for analysis in this study shows consistent relationships to variability of response for all four characteristics, nor are there significant differences in the variability of response for returns for men as compared with returns for women. Although individual responses in the two field visits show a relatively high variability for age and education, the summarized results for these items in the two enumerations in terms of frequency distributions and averages, are stable. This is not so, however, in the case of employment status.

Providing the enumerators measure up to a certain minimum standard, as the enumerators did for this survey, there is little correlation beyond that point between the individual's subjective rating and his performance as measured by variability of response. Variability of response is to be expected even under the best enumerating conditions, and this may reach surprisingly high proportions, as is indicated by the figures quoted above.

Source of Information.—There were 2686 Philadelphia households in the revisit sample, with a count of 8848 persons 16 years of age and over in the first enumeration and 8843 in the second.[4] There were 8519 persons for whom data could be matched for both visits, although all of the information analyzed in this study was not reported on both visits for each person enumerated twice.[5] The comparative data presented show differences for those of the total group of 8519 persons for whom the items were reported twice.

Selection of Enumerators.—The 36 enumerators employed in the revisit program were, on the whole, more experienced in schedule interviewing than the 197 enumerators who made the first visit. Both groups of enumerators, however, were given the same training with respect to the use of the schedule. Both groups were composed of high school or college graduates most of whom had had business, teaching, or social work experience. The information called for in the second visit referred to conditions as of the date of the first visit, and the second enumerators had no information about the household except the address.[6] An average (median) of 8.7 days elapsed between the two visits, and all revisits had been completed within 23 days.

---

[4] These counts included 324 persons 16 years of age and over who were included in the second enumeration but not in the first, and 329 included in the first and not in the second. This difference in the count of persons 16 years of age and over arises to some extent from actual changes in household composition, residence, or age between the two visits and from differences in the application of the "household" concept by the two groups of enumerators, but is primarily the result of inconsistencies in the reporting of respondents.

[5] For persons enumerated in both visits, employment status was reported for all persons, marital status for all but a few individuals, age for 99 per cent, and education for 95.8 per cent. It may be noted that there were more failures to report when a different respondent gave the information on the second visit.

In 63 per cent of the households, the same person answered the questions on both visits of the enumerators. In about three-fifths of these cases, the respondent was the housewife. Household heads or other relatives were the respondents in the remainder of the cases in which the same respondent gave the information on both visits.

TABLE I

PER CENT REPORTING DIFFERENT RESPONSE BY CITY DISTRICT COMBINATIONS

City district combinations	All respondents			Same respondent			Different respondent		
	Total	Different		Total	Different		Total	Different	
	Number	Number	Per cent	Number	Number	Per cent	Number	Number	Per cent
Marital status									
Total*	8515	141	1.7	5373	87	1.6	3142	54	1.7
Districts 1, 7, 8, 9, 10	4293	58	1.4	2800	36	1.3	1493	22	1.5
Districts 2, 3, 4, 5, 6†	4222	83	2.0	2573	51	2.0	1649	32	1.9
Age									
Total*	8472	1086	12.8	5349	552	10.3	3123	534	17.1
Districts 1, 7, 8, 9, 10	4262	525	12.3	2786	270	9.7	1476	255	17.3
Districts 2, 3, 4, 5, 6†	4210	561	13.3	2563	282	11.0	1647	279	16.9
Employment status									
Total*	8519	843	9.9	5375	455	8.5	3144	388	12.3
Districts 1, 7, 8, 9, 10	4293	367	8.5	2800	204	7.3	1493	163	10.9
Districts 2, 3, 4, 5, 6†	4226	476	11.3	2575	251	9.7	1651	225	13.6
Education									
Total*	8165	1911	23.4	5204	1133	21.8	2961	778	26.3
Districts 1, 7, 8, 9, 10	4092	907	22.2	2706	552	20.4	1386	355	25.6
Districts 2, 3, 4, 5, 6†	4073	1004	24.7	2498	581	23.3	1575	423	26.9

* Excludes persons not reporting items tabulated.

† Districts with 40 per cent or more Negro and/or foreign-born residents.

Method of Analysis.—The results of a matching of the two sets of schedules on a household basis, as distinct from the persons composing the households, have already been reported in another publication.[7] The present analysis is made on the basis of a posting of the two sets of responses for each person in the sample households. The data were tabulated by the sex of the person reported, by whether the respondent was the same or different, and by city district.[8] The proportion of individuals for whom different information was reported on the second as compared with the first visit varied only slightly as between districts in the city. Variability of response was measured by the degree to which the returns for the two visits fell into different class-intervals when grouped.[9]

---

[6] The 2686 addresses for the 1937 revisit program had been selected at random from a larger sample of 45,927 households in 149 selected School Blocks scattered throughout the city of Philadelphia, which had been used as a sample in a series of surveys of employment and unemployment from 1931 to 1938. The purpose of the check was to test the application of schedule concepts by the original enumerators as well as to check variability of response. The selection of revisit addresses was such that the work of all the original enumerators could be reviewed.

[7] This matching showed a relatively high variability of response for such items as age, employment status, occupation and industry, duration of jobs or unemployment periods, and years of residence in the city, as compared with such items as race, nativity, relationship to household head, and marital status. Margaret W. Bell and Gladys L. Palmer, *Employment and Unemployment in Philadelphia in 1936 and 1937*, Part 2, W.P.A. National Research Project in cooperation with the Industrial Research Department, University of Pennsylvania, 1938, Appendix A.

TABLE II

PER CENT REPORTING DIFFERENT RESPONSE BY NUMBER OF DAYS IN FIELD AT FIRST VISIT

Days in field at first visit	All respondents			Same respondent			Different respondent		
	Total	Different		Total	Different		Total	Different	
	Number	Number	Per cent	Number	Number	Per cent	Number	Number	Per cent
Marital status									
Total*	8484	140	1.7	5346	86	1.6	3138	54	1.7
First week	4631	77	1.7	2899	44	1.5	1732	33	1.9
Second week and later	3853	63	1.6	2447	42	1.7	1406	21	1.5
Age									
Total*	8442	1084	12.8	5323	550	10.3	3119	534	17.1
First week	4609	579	12.5	2890	298	10.3	1719	281	16.3
Second week and later	3833	505	13.2	2433	252	10.4	1400	253	18.1
Employment status									
Total*	8488	835	9.8	5348	448	8.4	3140	387	12.3
First week	4653	471	10.1	2901	248	8.6	1752	223	12.7
Second week and later	3835	364	9.5	2447	200	8.2	1388	164	11.8
Education									
Total*	8137	1905	23.4	5180	1127	21.8	2957	778	26.3
First week	4430	1053	23.8	2799	633	22.8	1631	420	25.7
Second week and later	3707	852	23.0	2381	494	20.7	1326	358	37.1

* Excludes persons not reporting items tabulated.

Factors in the Variability of Individual Response.—For all items except marital status, variability of response was significantly higher in cases in which the respondent was a different person on the two visits. The accompanying tables show this distinction. There seem to be no significant differences in the variability of response in reports for men as compared with women; consequently the data for the sexes have been combined in the analysis of differences in individual response.

Variability of response is found to be higher for districts with 40 per cent or more Negro or foreign-born residents or both combined, than for districts predominantly composed of native-born white residents

---

[8] The districting of the city follows the pattern of the School District divisions in use in 1937 when the original survey was made. All sections of the city were included.

[9] The class-intervals used for frequency distributions will be found in Table IV.

(Table I). The difference is not significant for returns for age, but is significant for returns for marital status, education, and employment status. It may be noted that this relationship holds for the returns for all four items when these were given by the same respondent, and for all items except age for returns given by a different respondent.

TABLE III

PER CENT REPORTING DIFFERENT RESPONSE BY DAYS BETWEEN TWO VISITS

Days between two visits	All respondents			Same respondent			Different respondent		
	Total	Different		Total	Different		Total	Different	
	Number	Number	Per cent	Number	Number	Per cent	Number	Number	Per cent
Marital status									
Total*	7394	123	1.7	4656	76	1.6	2738	47	1.7
1- 7	1544	25	1.6	976	15	1.5	568	10	1.8
8-14	5376	84	1.6	3417	54	1.6	1959	30	1.5
15-23	474	14	3.0	263	7	2.7	211	7	3.3
Age									
Total*	7354	944	12.8	4635	480	10.4	2719	464	17.1
1- 7	1533	197	12.9	971	89	9.2	562	108	19.2
8-14	5350	672	12.6	3402	359	10.6	1948	313	16.1
15-23	471	75	15.9	262	32	12.2	209	43	20.6
Employment status									
Total*	7395	735	9.9	4656	393	8.4	2739	342	12.5
1- 7	1544	150	9.7	976	68	7.0	568	82	14.4
8-14	5377	534	9.9	3417	303	8.9	1960	231	11.8
15-23	474	51	10.8	263	22	8.4	211	29	13.7
Education									
Total*	7085	1658	23.4	4509	976	21.6	2576	682	26.5
1- 7	1472	338	23.0	937	193	20.6	535	145	27.1
8-14	5159	1216	23.6	3319	728	21.9	1840	488	26.5
15-23	454	104	22.9	253	55	21.7	201	49	24.4

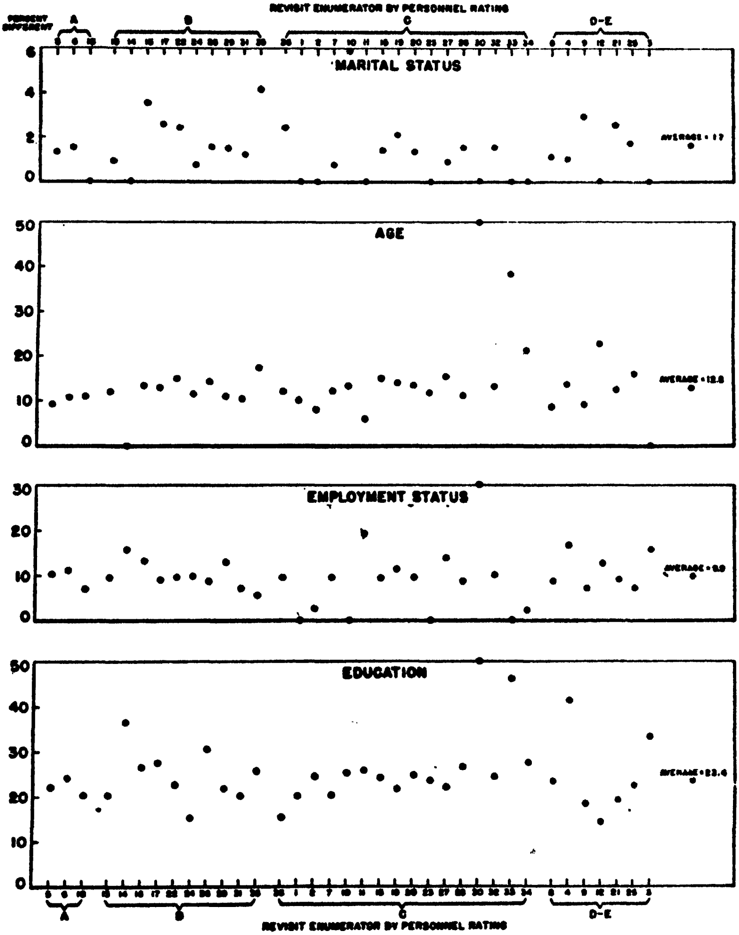
* Excludes persons not reporting items tabulated.

Compared with later returns, the results obtained during the first week of the first enumeration show only a slightly higher variability in the returns for marital status, employment status, and education, but not for age (Table II). It may be inferred that even well trained enumerators may become careless after the initial training wears off, and the effect of this carelessness almost counterbalances the steadying influence of experience. Moreover, the greater the lapse of time between the two visits, the greater the variability of response on marital status, age, and employment status (Table III), but not on education.

Variability Arising from Rating of Revisit Enumerators.—The revisit enumerators were rated subjectively A, B, C, D, E (footnote 3), and the accompanying chart was plotted to learn whether there is any relation between an individual's rating and variability of response, as

VARIATION IN RESPONSE FOR MARITAL STATUS, AGE, EMPLOYMENT STATUS, AND EDUCATION

By Personnel Rating of Revisit Enumerator



measured by the per cent of responses which differed in the two visits. There is a slight indication that A and B revisit enumerators showed less variability than those graded D and E. A few revisit enumerators showed variability outside the control limits (not drawn), but they are found in B ratings as well as D and E. This chart is consistent with the statement regarding the enumerators, that beyond a certain standard (surpassed by even the E revisit enumerators), there is little relation between performance as measured by variability of response and rating.

This does not mean that the general quality of schedule reporting is not better with experienced than with inexperienced enumerators, but that within any given framework of schedule concepts and training and qualifications of enumerators, variability of response is to be expected.

These findings support the evidence of another study in which variability of response in the work histories of skilled metal workers in Minneapolis was analyzed. In this study, many more factors were controlled than in a census type of survey, because a homogeneous group of workers was interviewed. Nevertheless, it was found that there was no consistent relationship between variability of response and the personality of interviewers or the lapse of time between the two interviews. Moreover, there was no consistent relationship between variability of response and a worker's age, education, or his nativity. The major factors which appeared to influence variability of response in this study of work history information were the simplicity or complexity of the data to be recalled and combinations of personal or psychological factors in response which could not be measured from available statistical data.[10]

Effects of Variation in Individual Response on the Summarized Results.—Although the degree of variability of response for individuals was relatively high for age and education, the effects of this are less significant when the data from the two enumerations are summarized (Table IV). The frequency distributions and averages secured for age and education show no significant differences, either for the total, or for men and women separately; nor are there significant differences in the proportions reporting various marital status categories in the two enumerations.

With regard to employment status, on the other hand, there are significant differences in the proportions reporting in certain categories in the two enumerations, when compared with each other and with the results of a combination of the two visits. The second visit yielded a larger labor force count than the first visit because it secured a considerably higher unemployment count. These differences are significant when related to the results of both visits combined. In the case of men, the count of the unemployed, and in the case of the women, the counts of the labor force, of non-workers, and of the unemployed differ significantly. This finding is corroborated by the results of other comparisons of employment status returns for the same individuals.[11]

The finding that, in general, summarized results are more stable than individual responses, is supported by other studies.[12] Many inconsist-

---

[10] Unpublished manuscript by Robert Dinkel in the files of the National Research Project of the Works Projects Administration.

[11] Palmer, *op. cit.*, footnote 2.

[12] Read Bain, "Stability in Questionnaire Response," *American Journal of Sociology*, November 1931, pp. 445–453; Ruth S. Cavan, "The Questionnaire in a Sociological Research Project," *American*

TABLE IV

MARITAL STATUS, AGE, EMPLOYMENT STATUS, AND EDUCATION BY SEX AS REPORTED IN THE FIRST AND SECOND ENUMERATIONS

	First enumeration						Second enumeration					
	Total		Men		Women		Total		Men		Women	
	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent
Marital Status												
Total*	8846	100.0	4449	100.0	4397	100.0	8822	100.0	4431	100.0	4391	100.0
Single	3184	36.0	1775	39.9	1409	32.1	3213	36.4	1793	40.5	1420	32.3
Married	4631	52.3	2335	52.5	2296	52.2	4535	51.4	2279	51.4	2256	51.4
Other†	1031	11.7	339	7.6	692	15.7	1074	12.2	359	8.1	715	16.3
Age												
Total*	8834	100.0	4450	100.0	4384	100.0	8782	100.0	4419	100.0	4363	100.0
16–19	970	11.0	465	10.4	505	11.5	987	11.3	471	10.7	516	11.8
20–24	1241	14.0	648	14.6	593	13.5	1227	14.0	648	14.7	579	13.3
25–29	1022	11.6	501	11.3	521	11.9	1011	11.5	479	10.8	532	12.2
30–34	790	8.9	417	9.4	373	8.5	761	8.7	401	9.1	360	8.2
35–39	818	9.3	389	8.7	429	9.8	827	9.4	412	9.3	415	9.5
40–44	834	9.4	422	9.5	412	9.4	819	9.3	395	8.9	424	9.7
45–49	785	8.9	402	9.0	383	8.8	782	8.9	393	8.9	389	8.9
50–54	739	8.4	381	8.6	358	8.2	732	8.3	385	8.7	347	8.0
55–59	545	6.2	289	6.5	256	5.8	530	6.0	287	6.5	243	5.6
60–64	451	5.1	210	4.7	241	5.5	463	5.3	223	5.0	240	5.5
65 and over	639	7.2	326	7.3	313	7.1	643	7.3	325	7.4	318	7.3
Median age		37.4		37.5		37.3		37.5		37.6		37.3
Employment Status												
Total*	8519	100.0	4266	100.0	4253	100.0	8519	100.0	4266	100.0	4253	100.0
Labor force	5444	63.9	3858	90.4	1586	37.3	5569	65.4	3891	91.2	1678	39.5
Employed	3312	38.9	2317	54.3	995	23.4	3222	37.8	2249	52.7	973	22.9
Unemployed	1797	21.1	1268	29.7	529	12.4	2017	23.7	1374	32.2	643	15.1
Emergency	335	3.9	273	6.4	62	1.5	330	3.9	268	6.3	62	1.5
Not seeking work	3075	36.1	408	9.6	2667	62.7	2950	34.6	375	8.8	2575	60.5
Education												
Total*	8442	100.0	4218	100.0	4224	100.0	8207	100.0	4069	100.0	4138	100.0
0	483	5.7	215	5.1	268	6.4	470	5.7	194	4.8	276	6.7
1– 4	747	8.9	374	8.9	373	8.8	714	8.7	370	9.1	344	8.3
5– 8	4206	49.8	2129	50.5	2077	49.2	4180	51.0	2121	52.1	2059	49.8
9–12	2651	31.4	1284	30.4	1367	32.4	2501	30.5	1176	28.9	1325	32.0
13–16	318	3.8	185	4.4	133	3.1	290	3.5	164	4.0	126	3.0
17 and over	37	0.4	31	0.7	6	0.1	52	0.6	44	1.1	8	0.2
Median		7.8		7.9		7.8		7.8		7.8		7.8

* Excludes persons not reporting items tabulated.

† Widowed, separated, divorced.

encies in reporting are inevitably compensating in character when the data are combined into frequency distributions. The extent to which variability in individual responses affects cross-classifications, such as employment status by age and sex or age and sex by education, has not been explored here but the chances of variability in the summarized results of such cross-tabulations are obviously greater.

It is clear from this study that a complete census of any group in the population is subject to some error or variability in reporting, regardless of the perfection of schedule concepts and definitions or the training of enumerators. It is not necessary to emphasize that variability of response affects the analysis of all population data, and is also important in the development of sampling techniques for enumerative studies. The choice of economic or social controls for stratification of sampling and sub-sampling units must be determined, in part, by the relative reliability of the information obtained in the survey used for controls.

Journal of Sociology, March 1933, pp. 721-727; Alberta Curtis, "The Reliability of a Report on Listening Habits," *Journal of Applied Psychology*, Volume 23, 1939, pp. 127-130; Mapheus Smith, "A Note on Stability in Questionnaire Response," *American Journal of Sociology*, March 1933, pp. 713-727.

TIME SERIES SIGNIFICANCE TESTS BASED ON SIGNS OF DIFFERENCES*

BY GEOFFREY H. MOORE
National Bureau of Economic Research
AND W. ALLEN WALLIS
Stanford University

RECENTLY we proposed a significance test, appropriate to time series, based on signs of first differences in either the observations or some derivative sequence.[1] The test is made by tabulating the frequency distribution of phase durations—in other words the frequency distribution of runs in signs of first differences—and comparing it with the distribution expected if the series were random, $\frac{5(N-3)}{12}$, $\frac{11(N-4)}{60}$ and $\frac{4N-21}{60}$ being the expected frequencies for phases of durations one, two, and over two, respectively. Since this expected distribution is the same for all continuous populations, such a test has a high degree of generality. The ideal statistic for the comparison is not known, but for practical purposes it suffices to calculate an ordinary χ^2 for goodness of fit (denoting it by χ_p^2) and refer 6/7 of its value to the usual χ^2 table for two degrees of freedom when $\chi_p^2<6.3$, or refer its whole value to a special tabulation of χ^2 for $2\frac{1}{2}$ degrees of freedom when $\chi_p^2>6.3$.

A variation upon the χ_p^2 test is to tabulate separate frequency distributions of durations of "expansions" (runs up) and "contractions" (runs down). Each distribution may be compared with an expected distribution having frequencies half as large as those expected for the combined distribution, or a test of homogeneity may be applied to the two distributions. Theoretical difficulties attach to either alternative, however.[2]

An alternative (but not independent) test can be made by simply counting the number of phases, i.e., the number of runs in signs of differences.[3] This is equivalent to counting the number of turning points, i.e., reversals in direction of movement. Unless N, the number of ob-

---

* This paper is a byproduct of work at the National Bureau of Economic Research, 1939–40, under Research Associateships provided by the Carnegie Corporation of New York.

[1] W. Allen Wallis and Geoffrey H. Moore, *A Significance Test for Time Series and Other Ordered Observations*. National Bureau of Economic Research *Technical Paper 1* (September, 1941), 59 pp. For a brief summary of this monograph, see W. Allen Wallis and Geoffrey H. Moore, "A Significance Test for Time Series Analysis," this JOURNAL, 36 (1941), pp. 401–9.

[2] See the *Technical Paper*, *op. cit.*, pp. 40–41. In addition to the difficulties there mentioned, it seems doubtful that the χ_p^2 distribution is appropriate to chi-squares derived from separate distributions of expansions and contractions.

[3] *Op. cit.*, Sec. VII.

servations, is small (in which case tabulations of the exact distributions are available), such counts are normally distributed with variance of $\frac{16N-29}{90}$, and mean of $\frac{2N-4}{3}$ for the number of turning points or $\frac{2N-7}{3}$ for the number of phases.

The present paper discusses three additional tests based on signs of differences:

(1) The total number of signs of each kind may be counted. The exact distributions are tabulated in Table I for $N \leqq 12$, and for longer series the number of signs of a given kind is normally distributed about $\frac{N-1}{2}$ with variance of $\frac{N+1}{12}$.

(2) The regularity of a periodic pattern, as judged by the tendency of the signs of one kind to cluster together in the pattern, can be tested, making allowance for any excess of signs of one kind. The exact probabilities can be calculated from a recursion formula, but have been tabulated only for the particular case of a nine point pattern.

(3) The existence of correlation between time series, abstracting from the effect of their trends, may be studied by making a 2×2 table of the signs of differences in the two series. The method may also be extended to more than two series. Although such a test would be of great value, the sampling distribution is known only for the special case of randomly arranged first differences.

In conclusion, a few general considerations relating to tests based on signs of differences are discussed.

I

A quick test of the randomness of a series can be made by counting the plus and minus signs in the first differences. A series of N observations has $N-1$ differences, and if the observations are in random order the expected number of plus or minus signs is obviously $\frac{N-1}{2}$. The variance of the number of signs of one kind is $\frac{N+1}{12}$, and the distribution rapidly approaches normality as N increases.

These results were reached by working with exact distributions for $N \leqq 12$, the exact distributions being calculated from the recursion relation

$$\phi_N(m) = (N - m)\phi_{N-1}(m - 1) + (m + 1)\phi_{N-1}(m),$$

where $\phi_N(m)$ represents the number of permutations of N different numbers that produce exactly m positive (or m negative) differences.[4] For a series of values of N the variances were found to have constant first differences of 1/12; since when $N=2$ the variance is $\frac{1}{4}$, the general expression[5] for the variance is $\frac{N+1}{12}$. The third moment about the mean is zero for all values of N, since the distributions are symmetrical. A series of values of the fourth moment about the mean shows, when $N \geqq 4$, constant second differences of 1/24; since when $N=4$ the fourth moment about the mean is 23/48 and the first difference 53/240, the general expression for the fourth moment about the mean is $\frac{5N^2+8N+3}{240}$.

As N increases, the kurtosis, measured by the ratio of the fourth moment about the mean to the squared variance, approaches 3, the value for a normal distribution. For all values of N the skewness is, of course, zero, as it is for a normal distribution. Comparison of the exact distribution for $N=12$ with that based on a normal approximation shows fair agreement; so for $N>12$ the normal distribution apparently provides a satisfactory approximation. When using a normal approximation the Laplace correction for continuity, which consists of subtracting $\frac{1}{2}$ from the discrepancy between observed and expected frequencies, should be applied. (A formula for calculating the normal deviate is given in the note to Table I.) For $N \leqq 12$ the exact distributions are shown in Table 1.

A test of the number of signs of one kind may serve as a test of correlation between two variables. The paired values are arranged in order according to one variable, and the signs of differences are tabulated for the other variables. Such a test is perhaps most appropriate for the detection of trend—the correlation of a series with time—for here no question arises as to which of the variables is to be arranged in order. In other cases the choice may not be obvious, and the result may depend upon which variable is selected.

---

[4] This relation was derived empirically from equation 4, page 19, of the *Technical Paper*. As is pointed out there (pages 5 and 11), a distribution of signs of differences is exactly the same for random arrangements of N different numbers as for samples of N drawn from any continuous population.

After this paper was in proof we discovered that exactly the same recursion relation given above was given in a different context by Major P. A. MacMahon, "Second Memoir on the Compositions of Numbers," *Philosophical Transactions of the Royal Society of London*, 207A (1908), pp. 65–134. The recursion relation given above appears in Art. 30, p. 83. At the end of Art. 33, p. 81 of the same memoir, MacMahon gives an explicit formula which, in our notation, may be written

$$\phi_N(m) = \sum_{i=0}^{m} (-1)^i \binom{N+1}{i} (m+1-i)^N.$$

He also tabulates $\phi_N(m)$ through $N=6$ on page 81.

[5] This expression may also be deduced easily from the familiar theorem that the variance of the sum of k variates is the sum of the k variances plus twice the sum of the $[k(k-1)]/2$ covariances.

TABLE I

EXACT PROBABILITY THAT THE DISCREPANCY BETWEEN OBSERVED AND EXPECTED NUMBERS OF MINUS SIGNS IN THE FIRST DIFFERENCES OF A RANDOM SERIES OF N OBSERVATIONS WILL EQUAL OR EXCEED (*IN ABSOLUTE VALUE*) THE DISCREPANCY CORRESPONDING WITH ANY OBSERVED NUMBER OF MINUS SIGNS, TWO TO TWELVE OBSERVATIONS; AND NORMAL APPROXIMATION FOR TWELVE OBSERVATIONS

m*	$N=2$	$N=3$	$N=4$	$N=5$	$N=6$	$N=7$	$N=8$	$N=9$	$N=10$	$N=11$	$N=12$ Exact	$N=12$ Approximate	m
0	1.000,000	0.333,333	0.083,333	0.016,667	0.002,778	0.000,397	0.000,050	0.000,006	0.000,001	0.000,000	0.000,000	0.000,00	0
1	1.000,000	1.000,000	1.000,000	0.450,000	0.161,111	0.048,016	0.012,302	0.002,772	0.000,559	0.000,102	0.000,017	0.000,12	1
2		0.333,333	1.000,000	1.000,000	1.000,000	0.520,635	0.225,248	0.083,284	0.026,926	0.007,750	0.002,014	0.003,95	2
3			0.083,333	0.450,000	1.000,000	1.000,000	1.000,000	0.569,582	0.277,803	0.118,154	0.044,551	0.054,67	3
4				0.016,667	0.161,111	0.520,635	1.000,000	1.000,000	1.000,000	0.606,074	0.321,454	0.336,67	4
5					0.002,778	0.048,016	0.225,248	0.569,582	1.000,000	1.000,000	1.000,000	1.000,00	5
6						0.000,397	0.012,302	0.083,284	0.277,803	0.606,074	1.000,000	1.000,00	6
7							0.000,050	0.002,772	0.026,926	0.118,154	0.321,454	0.336,67	7
8								0.000,006	0.000,559	0.007,750	0.044,551	0.054,67	8
9									0.000,001	0.000,102	0.002,014	0.003,95	9
10										0.000,000	0.000,017	0.000,12	10
11											0.000,000	0.000,00	11

* m is the number of minus signs.

NOTE: For larger values of N, $\pm \dfrac{\sqrt{3}(|2m-N+1|-1)}{\sqrt{N+1}}$ may be regarded as normally distributed with zero mean and unit variance.

As an illustration, this test of the sign counts has been applied to sweet potato production, yield per acre, and acreage harvested in the United States,[6] 1868–1937. The production series shows 37 rises and 32 declines. Since $N=70$ (observe that N is the number of observations, not differences), the expected number of rises or declines is 34.5, which differs from the observed number by 2.5, or 2.0 when corrected for continuity. The variance is 71/12, or the standard deviation 2.432. Taking $2.0/2.432=0.822$ as a normal deviate, we find a probability of 0.41 for so large a departure from expectation. Hence there is no evidence of trend. For acreage, the probability is about 0.0000000001, clearly indicating the presence of trend.[7] Actually, the charts of the two series suggest that approximately the same degree of trend is present in both series. The production series, however, contains random movements whose year-to-year variations are much greater than the yearly trend effects; whereas in the acreage series, particularly in the earlier years, the random variations are small relative to the trend. Any test of significance judges systematic variations by the relation of their magnitude to that of random variations; in this case the consequence is detecting the trend in the more stable acreage series and missing a similar trend in the more erratic production series. (This point is discussed further in the final section.) The yield series shows 35.5 rises and 33.5 falls (counting the one year of no change as half a unit in each category). Since as great a difference would result 84 per cent of the time from random causes alone, there is no evidence of trend; and this accords with the appearance of the chart.

It should be emphasized that the null hypothesis is that the *observations* are in random order, not that the first differences are random. To illustrate the distinction, suppose that of the six differences in a series of seven observations, 5 are of one sign and 1 of the other. On the null hypothesis that the seven numbers are in random order, we find the exact probability of a division as unequal as five-to-one to be 0.048 (Table I), so should probably conclude that the discrepancy is significant. Had we, on the other hand, formulated the null hypothesis that the *differences* are random and equally likely to be positive or negative we should have used a binomial distribution and found a probability of 0.219, more than 4½ times as great and definitely not significant. The variance on the null hypothesis that the differences are random and equally likely to be positive or negative is $(N-1)/4$, and therefore (when $N>2$) exceeds the variance on the null hypothesis that the observations are random; as N increases, the ratio between the two variances approaches 3. Hence the first hypothesis yields larger probabilities for given departures from expectation.

---

[6] The original data (taken from *Agricultural Statistics, 1939*, p. 243) and a chart are shown in the *Technical Paper*, pp. 37–39.

[7] Similar results were obtained in the *Technical Paper*, pp. 40–41, by testing separate frequency distributions of expansion and contraction phases. The number of signs of one kind is of course equivalent to the total duration of phases of one kind, hence is related to their frequency distribution by duration. The test of sign count or total duration, however, is free from the theoretical difficulties that attach to a test of the frequency distributions.

In economic research it is not always easy, but it is of course essential, to determine which of the two null hypotheses is appropriate. In testing the randomness of changes in stocks of commodities, for example, it might be appropriate to use a binomial distribution in connection with signs of first differences of stocks; but in testing whether stocks themselves constitute a random series, the appropriate distribution is the one given here. It may be observed that an inequality in the sign counts which cannot be accounted for on the assumption of random differences certainly cannot be accounted for on the assumption of random observations; and an inequality that can be explained by the hypothesis of random observations can be still more easily explained by the hypothesis of random differences.

II

A second kind of problem for which the signs of differences provide an appropriate significance test arises in periodogram analysis and similar procedures. Suppose a series has been divided into equal segments, or segments derived according to some external criterion, e.g., the turning dates of cycles in general business. It may be desired to test whether the movements within such segments tend to be smooth and regular rather than zig-zags, since regularity can normally be taken as indirect evidence that the period used is reflected in the data. Perhaps a description of the particular situation that resulted in this application will clarify the problem.

One feature of the method of studying cyclical behavior devised by the National Bureau of Economic Research involves dividing time series into segments corresponding with the dates of so-called reference cycles, i.e., cycles in general economic activity, each segment being bounded by two consecutive reference troughs.[8] Within each reference cycle, the series is expressed in reference cycle relatives; that is, each observation is divided by the average value of the series during that reference cycle. Each reference cycle is then divided into nine stages, and the mean of the relatives within each stage is computed. The first and ninth stages each consist of three months centered on the reference troughs and the fifth stage consists of three months centered on the reference peak; stages II, III, and IV correspond with successive thirds of the reference expansion, stages VI, VII, and VIII with successive thirds of the reference contraction. The nine means thus depict the pattern of behavior during a reference cycle.

---

[8] Cf. Wesley C. Mitchell and Arthur F. Burns, "The National Bureau's Measures of Cyclical Behavior," National Bureau of Economic Research *Bulletin 57* (July 1, 1935), 20 pp.; or better, their forthcoming volume, *Methods of Measuring Cyclical Behavior*.

Now if the series being analyzed contains cycles, any one reference pattern will be fairly regular, though its peaks and troughs need not conform to the reference dates. But an average reference cycle pattern made by averaging the reference stage standings over a number of reference cycles will tend to be simply a random sequence, unless the reference dates bear some relation to the behavior of the series—unless, in other words, the series conforms with general business activity.

The nine points of an average reference cycle pattern are rather few for a satisfactory test of the distribution of phase durations or of the total number of reversals. Furthermore, both these tests or one based on the total sign count would reflect the intra-cycle trend, which is retained in the reference cycle patterns. But from our recursion formula[9] it is possible to answer the question: Given the number of plus and minus signs in the average pattern, do they tend to cluster more than would be expected if they had occurred in a set of nine random standings?

To answer this question, consider the integers 1 to 8 arranged around a circle, with the signs of their first differences entered between them. The circular arrangement allows for the fact that stage IX in one cycle is based on the same data as stage I in the next, and eliminates the arbitrariness of taking up the cycle at any particular stage. Since 8 is the largest of the integers, the difference to one side of it must be positive and to the other side negative. From the recursion formula, the probability of each of the 64 arrangements of the other six signs that can occur among the 5,040 permutations of the digits 1 to 7 can be computed. Actually, we need consider only those of the 64 arrangements of signs that produce the observed count of pluses and minuses; if there are m minus signs ($1 \leqq m \leqq 7$), there are as many of these as there are combinations of six things taken $m-1$ at a time. The relevant arrangements of signs are then grouped according to the amount of clustering, as in Table II. The probability of as much clustering as is shown by each possible grouping is obtained by summing the probabilities of sequences of signs corresponding with each group.

This test is in a sense an amalgam of the test of phase durations and the test of turning points. Except for the circular arrangement, it is like the test of turning points in that the fewer the turning points the greater the clustering. But for a given number of turning points, the de-

---

[9] *Technical Paper*, Equation 4, p. 19.

TABLE II

EXACT PROBABILITY OF AS MUCH (OR MORE) REGULARITY OF PATTERN FROM EIGHT OBSERVATIONS ARRANGED AT RANDOM IN A CLOSED SEQUENCE AS IS SHOWN BY EACH POSSIBLE GROUPING OF SIGNS OF DIFFERENCES

4:4 Ratio between signs		5:3 Ratio between signs		6:2 Ratio between signs	
Grouping*	Probability	Grouping*	Probability	Grouping*	Probability
4 4	.0083	5 3	.0126	6 2	.0500
3–1 3–1	.0993	4–1 2–1	.1453	5–1 1–1	.2667
3–1 2–2	.2368	3–2 2–1	.3955	4–2 1–1	.7167
2–2 2–2	.2914	3–1–1 1–1–1	.6448	3–3 1–1	1.0000
2–1–1 2–1–1	.8874	2–2–1 1–1–1	1.0000		
1–1–1–1 1–1–1–1	1.0000				

* Entries under "grouping" indicate in the first line the lengths of runs in one sign, in the second line the lengths of runs in the other sign. Thus $\begin{matrix}3\text{–}1\\2\text{–}2\end{matrix}$ indicates that the pluses (minuses) are contained in one group of 3 and one group of 1 and the minuses (pluses) in two groups of 2 each.

gree of clustering is greater when the signs of one kind fall mostly in a single group (phase) than when they are more evenly distributed among the groups.

III

The test of cyclical conformity just described is based simply upon the regularity of the average pattern. Some types of systematic movement, e.g., seasonal, do not imply this type of regularity—indeed, it is possible that a series might follow business cycles systematically, but in an irregular fashion. This suggests the need for another kind of test, one to determine, for example, whether movements during corresponding stages of successive reference cycles are consistent. Such a test would be one of a general class of correlation tests based on signs of first differences.

Suppose it is desired to test whether there is a relation between the nine-point patterns of two reference cycles, the seasonal patterns given by twelve monthly figures for each of two years, or any two time series of equal length. A 2×2 table can be constructed by listing the signs of differences for each series and tabulating the number of instances in which (a) both series rise, (b) the first rises and the second falls, (c) the first falls and the second rises, and (d) both fall. A homogeneity test of this double dichotomy, treating the marginal totals as fixed, would have the advantages of simplicity, generality, and elimination of at least the primary effects of trend.

The usual procedures for analyzing double dichotomies are based on the assumption that the entries are random and independent. If each series is a random sequence, the differences are negatively correlated serially, and this means that a chance coincidence in direction of movement during one interval tends to produce another coincidence during the next interval. Furthermore, suppose the first series has an excess of positive differences; these are more likely to be located at certain parts of the series than at others.[10] If the second series also has more plus than minus signs they tend to fall in the same positions, thus producing coincidences even though the series are unrelated; but if the second series has more minus than plus signs, the minuses tend to fall where the pluses do for the first series, thus producing a spurious appearance of a negative relation between the two series.

A few tabulations that we have made of the exact distribution on the assumption that the observations are random suggest that the usual

---

[10] This may seem surprising since each of the $N-1$ differences in a random series is equally likely to be positive or negative. When it is known, however, that one sign has occurred less frequently than the other, the occurrences of the less frequent sign are more likely to be at the center than at the extremes of the series. This may be verified by the recursion formula referred to in footnote 9, but the following example illustrates the point: Suppose that of the $N-1$ differences in a series of N, exactly one is negative. If this negative difference is the mth in the series of $N-1$ differences, it means that (1) the first m observations are arranged in ascending order, (2) the last $N-m$ observations are arranged in ascending order, and (3) the greatest observation among the first m exceeds the least among the last $N-m$. Any partition of the N observations into two groups, of m and $N-m$ respectively, provides for one (and only one) fulfillment of the first two conditions; so there are

$$\frac{N!}{m!(N-m)!}$$

ways to satisfy conditions 1 and 2. Of these, all but one—that in which each observation among the $N-m$ exceeds each among the m—also satisfy the third requirement. Hence there are among the permutations of N different numbers exactly

$$\frac{N!}{m!(N-m)!}-1$$

which have all differences positive except the mth. This quantity summed over the range $m=1$ to $m=N-1$, or 2^N-N-1, is the number of permutations of N different quantities that produce exactly one negative sign among the signs of differences. So if a random sequence of N numbers has produced only one negative first difference (the same reasoning applies for one positive difference), the probability that it is the mth difference is

$$\frac{N! - m!(N-m)!}{m!(N-m)!(2^N-N-1)}$$

which has a maximum at $m=N/2$ when N is even or maxima at $m=(N-1)/2$ and $m=(N+1)/2$ when N is odd, and minima at $m=1$ and $m=N-1$. When there is more than one minus sign, but fewer than $(N-1)/2$, the situation is more complex. For example, two minus signs if adjacent are most likely to be in the center but if non-adjacent are most likely to be about one-third of the way from each end; they are more likely to be non-adjacent than to be adjacent and the net result appears to be a distribution with three modes, one at the center and the other two a little more than a third of the way from each end. As the number of minus signs rises toward $(N-1)/2$ the probability of a minus in any given position becomes more nearly uniform for all positions.

treatment of double dichotomies, though based on the assumption of random differences, may provide usable approximations unless N is very small or the number of plus and minus signs within either series is extremely unequal. When $N=8$ and each series has the same number of signs of a given kind, the comparison is as shown in Table III.

TABLE III

PROBABILITIES OF VARIOUS NUMBERS OF COINCIDENCES IN POSITION FOR SIGNS OF DIFFERENCES IN TWO RANDOM ARRANGEMENTS OF EIGHT OBSERVATIONS, EACH CONTAINING m POSITIVE (OR m NEGATIVE) DIFFERENCES; AND COMPARISON WITH APPROXIMATION BASED ON ASSUMPTION THAT DIFFERENCES ARE RANDOM AND INDEPENDENT

m*	Coincidences†	Probability	
		Exact	Approximate‡
1	1	0.202708	0.142857
	0	0.797292	0.857143
2	2	0.070409	0.047619
	1	0.470286	0.476190
	0	0.459305	0.476190
3	3	0.042049	0.028571
	2	0.342878	0.342857
	1	0.480747	0.514286
	0	0.134325	0.114286

* m is the frequency of the less frequent sign.

† Number of coincidences in position for the less frequent sign. The table also covers pairs of series in which one has m signs of one kind and the other has m signs of the opposite kind, if coincidences are counted for the two sets of m; in such cases inverse relationship is tested.

‡ Calculated by the method given in R. A. Fisher, *Statistical Methods for Research Workers*, Fifth and Later Editions (London and Edinburgh, 1934–40), Section 21.02.

It may be remarked that if the two dichotomies appear to be homogeneous when treated by the usual procedure, which implicitly assumes random differences, they surely would appear so if the proper null hypothesis of random observations were used; but nothing more can be said definitely until the sampling distribution appropriate to the assumption of random observations is known.

It is easy to extend this method of testing correlation to more than two series. For example, if it is desired to test the consistency of reference cycle behavior over several reference cycles, the signs of the differences for the eight stage-to-stage movements in the separate reference cycle patterns may be entered in a table of eight columns and as many rows as there are cycles. By counting the number of plus signs in each column, and comparing the counts with those expected if the total number of plus signs in the entire table were equally divided among the columns, we can see whether the movements during each interval tend to be consistent from cycle to cycle, and at the same time allow for the effect of trend on the total number of plus signs. In testing the significance of this tendency, however, difficulties with respect to the null hypothesis again rise. If the differences can be assumed random sequentially, the variance of the column counts is $\frac{-}{C}\left(1-\frac{-}{CR}\right)$ (where R is the number of rows, C the number of columns, and P the number of pluses in the table), and the ratio to this of the sum of squared differences between observed and expected column counts is distributed as χ^2 with $C-1$ degrees of freedom, provided R is not too small and the proportion of plus signs is not too far from one-half (when R is small or the number of pluses and minuses extremely unequal, exact distributions may be computed without great difficulty). If the appropriate null hypothesis is that the observations are sequentially random, we know of no valid procedure except the laborious one of tabulating exact distributions, unless the procedure assuming randomly arranged differences turns out to be satisfactory also for randomly arranged observations.

The desirability of a solution to this problem is enhanced by the possibility of using the same criterion, signs of differences, to test the randomness of the separate series. On the other hand, the technique would be restricted in scope by the fact that in many time series problems neither the hypothesis of sequential randomness in the observations nor that of sequential randomness in the differences is tenable. For example, in testing the correlation between two economic series we should allow for the presence of cyclical characteristics which tend to duplicate, or to cancel, any chance coincidences. (In these cases, methods suggested earlier may be preferable—arranging the pairs in order according to one variate, and counting the number of signs of one kind, the number of reversals, or the number of phases of each duration in the other variate.)

IV

The generality and simplicity of tests based on signs of differences is due to the fact that they utilize only information that is inherent in every sample. When there exists relevant information other than that contained in the sample, as for example the knowledge that the observations are from a normal or approximately normal population, the tests are "inefficient" in the sense that they fail to utilize the supplementary information. Where no such additional knowledge exists, however, they are more effective than tests of narrower scope, in that they do extract some information from samples to which the less general tests cannot be applied at all. Furthermore, even when some supplementary information is available, a test based on signs of differences may be quite efficient. For example, in certain types of data the signs of differences are more accurate than the magnitudes of either the observations or the differences.[11] Again, the problem studied may be such that direction of movement *per se* is significant and the magnitude of movements is irrelevant. A similar situation arises when it is desired to abstract from the effects of trend, for a test based on signs of differences is relatively insensitive to movements that are gradual in relation to the other movements in a series.

The greater the random movements in a series, the more likely are tests based on signs of differences to overlook systematic movements. This is true of any significance test, for all significance tests judge the systematic tendencies by comparison with random fluctuations. The weakness of tests based on signs of differences is, however, that the random variation ordinarily is reduced only slightly by increasing the amount of data; that is, the "averaging out" process is not very effective. It is possible to conceive of cases where increasing the length of a series would be highly effective (for example, in data having a trend rise of ten units per year, if the trend-adjusted data fall within a range of one unit), and also of cases where it would have absolutely no effect (for example, in data having a trend rise of one unit per year, if the trend-adjusted data never differ by less than ten units). The series on sweet potato acreage discussed previously lies toward the first, that on sweet potato production toward the second, of these extremes.

No general statement about the efficiency or inefficiency of tests based on signs of differences can be valid except in the particular context of a specific problem and body of data.

---

[11] With economic measures of the kind for which index numbers are ordinarily used, it may be certain that a change has been in a given direction (e.g., when all components of the index change in the same way), questionable how much the change has been (because of ambiguities in the weighting system), and meaningless to state an absolute standing. Similarly, the short-time directions of change of an index may be more meaningful than the long-time directions of change.

AN APPROACH TO THE MEASUREMENT OF FARM POPULATION PRESSURE IN WISCONSIN

BY THOMAS C. McCORMICK
University of Wisconsin

THE CONCEPT of an optimum population is actually so basic in current population theory that without it the latter would have little meaning or motive. It is, however, still essentially a philosophical abstraction that is sometimes discounted or even denied. In spite of its important social consequences, the idea has never been the subject of serious quantitative testing and study in concrete situations. The present paper reports the results of a quantitative approach to one aspect of the problem recently made in Wisconsin, which may be of interest as a methodological venture.

From the agricultural point of view, an optimum population implies that the per capita income of a farm population is in part a function of the ratio of the size of the population to the amount of land farmed. Accordingly, if interfering factors can be controlled, variations in the acreage of farm land per capita of the farm population from county[1] to county should be accompanied by differences in the per capita farm income, which should range from a minimum to a maximum.

Only data published by the state and federal governments were available. Since as a rule such large-scale data can come only from government sources, this was regarded as one of the conditions of the problem.

The search for an index of the income of the farm population was able to discover nothing better than a detailed and careful estimate of gross farm income by counties made by the Wisconsin Department of Agriculture and Markets in 1931 (no later estimates have been prepared). These county estimates were deflated and reduced to a per capita basis of the rural farm populations[2] in 1930. Although the index is not without faults, it was found to be rather highly correlated[3] with other indexes of agricultural prosperity, such as the value of crops and livestock products per capita of the rural-farm population in 1930 ($r^2=.81$), the percentage of farm homes with automobile, radio, and telephone ($r^2=.62$), and the per capital value of farmers' dwellings

---

[1] The county was the smallest unit for which data were published. Data were lacking for a time series analysis.

[2] *Wisconsin Agriculture*, Bulletin No. 140, p. 45, 1932; and 1930 Census of the United States, *Population*, Vol. III, Pt. 2, p. 1328. The gross farm income is the estimated value of all products grown. This was divided by 146 and multiplied by 100 to give the deflated index. The index of prices paid by farmers for commodities bought in 1930 is 146, with 1910–14 as the base period. See *Wisconsin Agriculture*, *op. cit.*, p. 10, column 20.

[3] Computed from a test sample.

($r^2=.64$). This index, because it represents the dependent variable, is represented by the symbol X_0, and will be referred to as "per capita income." The independent variable, the acreage of farm land per capita of the rural-farm population, was of course readily obtained for each county by dividing the total farm acreage by the rural-farm population[4] in 1930. It has the symbol X_1, and will be called the "land-man ratio." In studying the relationship between the per capita income (X_0) and the land-man ratio (X_1), the chief difficulty lay in the control of interfering factors. As usual in social research, this problem was only partially solved by holding constant farm land values per acre[5] (X_2) and type of farming[6] (X_3).

GENERAL METHOD AND RESULTS

The data and method of analysis are shown in Tables I through VI. Omitting the two highly urbanized counties of Milwaukee and Kenosha, the remaining sixty-nine counties in Wisconsin are divided into three broad types of farming areas, each area occupying two tables to itself. In Tables I, III, and V the counties in each area are classified first by mean value of farm land per acre (X_2), and second by land-man ratio (X_1). In any land-value class (row) in an area table (Tables I, III, and V), both type of farming and land value per acre are held roughly constant, so that variations in mean per capita income ($\overline{X}_0$) from one land-man ratio class (column) to another can probably be attributed to differences in the land-man ratio (X_1) and (unknown) associated factors. It can be seen by inspection that in every row (of the odd-numbered tables) the mean per capita income (X_0) increases as the land-man ratio (X_1) increases. The actual amount of this relationship, with land value (X_2) and type of farming (X_3) constant, is roughly measured by the squared partial correlation ratio, $\eta_{01.23}^2$, calculated for each row and averaged over the table. The $\eta_{01.23}^2$ values for Tables I, III, and V are .63, .68, and .49, respectively.

On the strength of the relationship just demonstrated between per capita income (X_0) and the land-man ratio (X_1), it is assumed that those counties that have the highest land-man ratio in a given row represent, from the point of view of per capita farm income, the most favorable adjustment of farm population to farm land in their class; and the mean land-man ratio of these "top" counties is taken as the criterion for adjusting upward the land-man ratios of the other counties in the same class (row). The farm population required to raise the land-man ratio of a county to that of the criterion counties in the same row is found by dividing the farm acreage of the county in question by the mean land-man ratio of the criterion counties.[7] This is called the "adjusted" farm population. A county index of relative population pressure in agriculture, symbolized by the letter *I*, is then obtained[8] by dividing the adjusted farm population by the census rural farm population in 1930, multiplying by 100, and subtracting from 200. (See column *I*, Tables II, IV and VI.) Thus, an index of 120 means that the rural farm population of a county must be reduced by 20 per cent in order to produce a land-man ratio equal to that of the criterion counties in the same land-value class and type of farming area, which of course have an index of 100.

---

[4] 1930 Federal Census, *Agriculture*, Vol. II, Pt. 1, p. 736.

[5] 1930 Federal Census, *Agriculture*, Vol. II, Pt. 1, pp. 736, 748. Deflated by dividing value per acre by 115 and multiplying by 100. The index of farm real estate values in 1930 is 115, with 1910–14 as base period. See *Wisconsin Agriculture*, *op. cit.*, p. 10, column 22.

[6] Type of farming areas were taken with modifications from P. E. McNall and W. J. Roth, *Forces Affecting Wisconsin Agriculture, with Resulting Types of Farming*, Research Bulletin 131, Agricultural Experiment Station of the University of Wisconsin, November, 1935.

TABLE I

TYPE OF FARMING AREA I, TWENTY-TWO WISCONSIN COUNTIES: METHOD OF ANALYSIS*

X_3	X_1	18–23	24–29	30–33	Totals
30–44	n	3	0	2	5
	$\bar{X}_0$	214.3		246.5	227.2
	ΣX_0	643		493	1136
	ΣX_0^2	143,605		121,549	265,154
	$(\Sigma X_0)^2/n$	137,816		121,525	258,099
45–59	n	8	4	1	13
	$\bar{X}_0$	214.9	303.0	328.0	250.7
	ΣX_0	1,719	1,212	328	3,259
	ΣX_0^2	376,641	369,882	107,584	854,107
	$(\Sigma X_0)^2/n$	369,370	367,236	107,584	817,006
60–84	n	4	0	0	4
	$\bar{X}_0$	252.8			252.8
	ΣX_0	1,011			1,011
	ΣX_0^2	256,677			256,677
	$(\Sigma X_0)^2/n$	255,530			255,530
Totals	n	15	4	3	22
	$\bar{X}_0$	224.9	303.0	273.7	245.7
	ΣX_0	3,373	1,212	821	5,406
	ΣX_0^2	776,923	369,882	229,133	1,375,938
	$(\Sigma X_0)^2/n_c$	758,475	367,236	224,680	1,350,391
	$\Sigma^r[(\Sigma X_0)^2/n_r]$				1,330,635
	$\Sigma^{rc}[(\Sigma X_0)^2/n_{rc}]$	762,716	367,236	229,109	1,359,061
	$(\Sigma X_0)^2/N$				1,328,402

* For meaning of symbols, see text.

NOTE: Tables I, III, and V would be somewhat more exact if the calculations had been weighted by the rural farm populations of the counties. In the case of estimates as rough as these, however, the extra labor required did not seem justified.

---

[7] Or, the adjusted farm population may be obtained by multiplying the observed rural farm population of the county by its index of population pressure, *I*, as described in the text, subtracted from 200 and expressed as a decimal.

[8] Or, the land-man ratio of the county may be divided by the criterion land-man ratio and the quotient subtracted from 2 and multiplied by 100.

TABLE II
TYPE OF FARMING AREA I: OBSERVED DATA AND ESTIMATES*

No.	County	X_0	X_1	X_2	P	$F.I.$	I	X_0'	P'	$F.'I.' = X_0' \times P'$
1	Jefferson	276	28.8	44	13,604	3,759,863	126	302	10,067	3,040,234
2	Outagamie	190	18.8	44	17,843	3,398,904	141	231	10,527	2,431,737
3	Shawano	177	22.6	32	19,200	3,394,589	129	206	13,632	2,808,192
4	Columbia	250	32.1	33	13,876	3,469,247	100	250	13,876	3,469,247
5	Green Lake	243	33.0	38	6,248	1,516,849	100	243	6,248	1,516,849
6	Brown	184	17.9	49	16,534	3,035,274	134	311	10,912	3,393,633
7	Door	171	23.0	47	10,043	1,720,137	115	227	8,537	1,937,899
8	Fond du Lac	251	23.1	52	18,737	4,695,548	115	306	15,926	4,873,356
9	Kewaunee	188	21.6	51	9,601	1,807,808	120	264	7,681	2,027,784
10	Manitowoc	207	19.7	47	17,791	3,685,205	127	309	12,987	4,012,983
11	Sheboygan	235	18.7	56	15,980	3,748,288	131	351	11,026	3,870,126
12	Washington	257	20.2	51	12,903	3,312,192	125	352	9,677	3,406,304
13	Winnebago	226	21.1	50	12,004	2,712,671	122	309	9,363	2,893,167
14	Dane	273	24.7	54	28,763	7,856,644	100	273	28,763	7,856,644
15	Dodge	299	24.4	54	21,112	6,316,027	100	299	21,112	6,316,027
16	Rock	296	27.1	47	15,442	4,570.137	100	296	15,442	4,570,137
17	Walworth	344	27.5	50	11,380	3,919,726	100	344	11,380	3,919,726
18	Green	328	32.2	50	11,077	3,628,767	100	328	11,077	3,628,767
19	Calumet	229	19.5	61	9,864	2,259,041	100	229	9,864	2,259,041
20	Ozaukee	246	17.9	75	7,386	1,818,973	100	246	7,386	1,818,973
21	Racine	262	18.6	80	9,419	2,466,507	100	262	9,419	2,466,507
22	Waukesha	274	19.5	75	15,222	4,175,274	100	274	15,222	4,175,274
Totals					314,078	77,267,671			270,124	76,692,607
Means		246†	23.0	52			114†	284†		

* Meaning of symbols used as column headings:

X_0 is deflated gross farm income per capita of the rural farm population, 1931.

X_1 is acres of farm land per capita of the rural farm population, 1930, referred to as land-man ratio.

X_2 is the deflated value of farm land per acre, 1930.

P is the rural farm population, 1930.

$F.I.$ is the total deflated gross farm income, 1931.

I is the index of relative farm population pressure.

X_0' is the adjusted deflated gross farm income per capita of the rural farm population, 1931.

P' is the adjusted rural farm population, 1931.

$F'I'$. is the adjusted total deflated gross farm income, 1931.

† Weighted by the rural farm populations of the counties.

By this method absolute optima populations are avoided and relative optima, established by the counties themselves, are used that change as the value of farm land changes.

Next, a rough estimate is made of the probable effect of the increased land-man ratios on the farm income per capita in each adjusted county and on the total farm income in the type of farming area. The results show that while increasing the number of acres per capita always increases per capita income, in each of the three types of farming areas it reduces the total farm income because the returns from a fixed amount of farm land decrease as cultivation becomes less intensive (cf. the *F.I.* and *F.'I.'* columns of Tables II, IV, and VI). This appears again in the fact that in Area II, for example, the coefficient of linear correlation between the land-man ratio and farm income *per acre* is $r = -.68$, or $r^2 = .46$. The operation of this principle is familiar in the thickly populated countries of Europe and Asia, where returns per acre are high but returns per worker are very low.

Finally, the investigation shows that in each type of farming area in Wisconsin relative over-population arises in greater degree from too many farm families than from too large farm families. There is a slight tendency for farm families to be larger in counties where the farms are smaller and incomes lower.

The steps in this analysis will be shown in some detail for type of farming Area I. In the cases of the other two areas only the end results will be given, except where certain points need explanation.

TYPE OF FARMING AREA I

Type of farming Area I, represented by Tables I and II, includes twenty-two counties located in the southern and eastern parts of the state. The dominant farm enterprise is dairying. The region is distinguished by the highest gross farm income per capita of the rural farm population, the weighted mean[9] being $246 (column X_0, Table II). The area also has the highest priced farm land, with a range of $32 to $80 and a mean of $52 per acre (column X_2, Table II). The mean land-man ratio, however, is only 23 acres of farm land per capita of the rural farm population, which is smaller than that of either of the other two areas (column X_1, Table II).

The relationship between per capita income (X_0) and land-man ratio (X_1) in this area, with type of farming (X_3) and value of land (X_2) constant, may be found from the formula

$$\eta^2_{01.23} = \frac{\sum^{r} \sum^{c} \frac{\left(\sum^{rc} X_0\right)^2}{n_{rc}} - \sum^{r} \frac{\left(\sum^{c} X_0\right)^2}{n_r}}{\sum^{r} \sum^{c} X_0^2 - \sum^{r} \frac{\left(\sum^{c} X_0\right)^2}{n_r}} \qquad (1)$$

where r refers to row, c to column, rc to cell, and n to frequency. In using this formula it is well to omit (from both numerator and denominator) any row that has less than two cells containing two or more X_0 entries each. Substituting from Table I,[10]

---

[9] Unless otherwise stated, all area means are unweighted averages of the county means. Weighting was usually found to make no practical difference in the results.

[10] It seemed advisable to avoid Bessel's correction and the correction for broad grouping. The assumptions of the latter make it invalid in this study, and since the two corrections tend to offset one another, it would be misleading to use the former alone. If both were made in the usual way, the values of η^2 would be materially raised here.

$$\eta^2_{01.23} \quad \frac{1,359,061 - 1,330,635}{1,375,938 - 1,330,635} \quad .63.$$

That is, roughly 63 per cent of the variance (standard deviation squared) in per capita income from county to county may be attributed to differences in the land-man ratio and associated factors, with the influence of differences in land values and type of farming removed.

It is also possible to estimate the extent to which per capita farm income reflects the combined influence of the two factors, land-man ratio, X_1, and land value per acre, X_2, with type of farming, X_3, held constant. For this purpose the following formula may be devised,

$$H^2_{0.12(3)} = 1 - \frac{\sum^{r}\sum^{c} X_0^2 - \sum^{r}\sum^{c} \frac{\left(\sum^{rc} X_0\right)^2}{n_{rc}}}{\sum^{r}\sum^{c} X_0^2 - \frac{\left(\sum^{N} X_0\right)^2}{N}} \tag{2}$$

where N is the total frequency of the table. In making such an estimate, cells that contain less than two entries should be omitted entirely. The value here is

$$H^2_{0.12(3)} = 1 - \frac{1,375,938 - 1,359,061}{1,375,938 - 107,584 - \frac{(5406 - 328)^2}{21}} = .58.$$

But $H^2_{0.12(3)}$ should be greater than $\eta^2_{01.23}$ if the factors X_1 and X_2 are held actually constant in the cells. As the X_1 and X_2 categories in Table I become narrower the value of $H^2_{0.12(3)}$ increases. To correct $H^2_{0.12(3)}$ for broad categories we may find $\eta^2_{02.3}$ for the whole table

$$\eta^2_{02.3} = \frac{\sum^{r} \frac{\left(\sum^{c} X_0\right)^2}{n_r} - \frac{\left(\sum^{N} X_0\right)^2}{N}}{\sum^{r}\sum^{c} X_0^2 - \frac{\left(\sum^{N} X_0\right)^2}{N}} \tag{3}$$

$$\eta^2_{02.3} = \frac{1,330,635 - 1,328,402}{1,375,938 - 1,328,402} = .05$$

and then use the formula

$$H^2_{0.12(3)} = \eta^2_{01.23}(1 - \eta^2_{02.3}) + \eta^2_{02.3} \tag{4}$$

$$H^2_{0.12(3)} = .63(.95) + .05 = .65.$$

We therefore assume that something like 65 per cent of the variance in per capita income between the counties is attributable to the combined influence of the land-man ratio and land values, leaving 35 per cent unexplained.

By a formula similar to (1),

$$\eta^2_{02.13} = \frac{\sum^{r}\sum^{c} \frac{\left(\sum^{rc} X_0\right)^2}{n_{rc}} - \sum^{c} \frac{\left(\sum^{r} X_0\right)^2}{n_c}}{\sum^{r}\sum^{c} X_0^2 - \sum^{c} \frac{\left(\sum^{r} X_0\right)^2}{n_c}} \tag{5}$$

$$\eta^2_{02.13} = \frac{1{,}359{,}061 - 1{,}350{,}391}{1{,}375{,}938 - 1{,}350{,}391} = .34.$$

A comparison of $\eta^2_{02.13} = .34$ with $\eta^2_{01.23} = .63$ shows that the relationship between the land-man ratio and per capita income is much closer than that between the per capita income and the value of farm land per acre.

In Table I it is seen that five counties have farm land with mean values ranging between \$30 and \$45 an acre. Three of the counties, Shawano, Outagamie, and Jefferson, have a mean land-man ratio of 21.7 acres and a mean per capita income of \$214. The other two counties, Green Lake and Columbia, have a mean land-man ratio of 32.6 acres and a per capita income of \$247. If, in view of our correlation analysis above, we attribute the difference of about \$32 in per capita income between these two groups of counties to the difference in their land-man ratios (and unknown associated factors), we may estimate that each increase of one acre in land-man ratio is accompanied by a mean gain of roughly \$3.13 in per capita income $\left(\frac{246.5 - 214.3}{32.0 - 21.7}\right)$. To raise the mean land-man ratio of the three "low" counties to 32 acres, we divide 32 into the land-man ratio of each of these counties, i.e., into 22.6 in the case of Shawano, into 18.8 in the case of Outagamie, and into 23.8 in the case of Jefferson, and subtract the resulting percentage from 200 in each case. We then obtain as the county indexes of relative farm population pressure in this land-value class: Jefferson 126, Outagamie 141, Shawano 129 (column I, Table II). That is, the rural farm population of Jefferson county would have to be reduced 26 per cent in order to provide 32 acres of farm land for the support of each individual member. Columbia and Green Lake, the criterion counties in this group, each has a population pressure index of 100, and their farm populations are left unchanged. When the farm populations of Jefferson, Shawano, and Outagamie counties are reduced in this way, the adjusted farm population for each of the three counties may be found by multiplying its 1930 rural farm population by the difference between 200 and the county's index of population pressure, expressed as a decimal. For example, for Jefferson county, with a rural farm population of 13,604 in 1930, the adjusted population is $13{,}604 \times (2.00 - 1.26) = 10{,}067$. For the other two counties the 1930 rural farm populations and their adjusted values are: Outagamie, 17,843 reduced to 10,527; Shawano, 19,200 reduced to 13,632 (column P', Table II).

If such adjustments in the land-man ratios of Jefferson, Outagamie, and Shawano counties were made, we infer that each county would probably have its per capita farm income increased by an amount equal roughly to the increase in its land-man ratio multiplied by the figure \$3.13 obtained above. Thus, in the case of Jefferson county, the land-man ratio is increased from 23.8 acres to 32 acres, or 8.2 acres, so that the gain in its per capita income is $8.2 \times 3.13 = \$25.67$. Since the observed per capita income of Jefferson is \$276, its adjusted per capita income is $276 + 25.67 = \$301.67$, or about \$302. Similarly, in Outagamie county the per capita income is raised from \$190 to \$231, and in Shawano county from \$177 to \$206 (column X_0', Table II).

By repeating these simple calculations for each land-value class in Table I, Table II (to the right of the upright double ruling) was worked out for the twenty-two counties in Area I. The land-man ratio of each of the counties numbered 6 through 13 in Table II was raised to 27 acres, and the addition of each acre was estimated to increase the per capita income by \$14.00.

The weighted index of farm population pressure for the whole of Area I is 114 (Table II). It is found by dividing the total observed rural farm population (P) into the total adjusted farm population (P'), and subtracting from 200, or $200 - \frac{270{,}124}{314{,}078} = 114$ (columns P and P' in Table II). That is, a reduction of 14 per cent in the rural farm population of this prosperous dairy section in southern and eastern Wisconsin is needed to raise the land-man ratios in all counties to a par with the land-man ratios observed in the counties with the highest land-man ratios in their respective land-value classes. A population of about 44,000 (i.e., 314,078 × .14) would be released from the farms, chiefly through migration to the cities, partly through birth control. As a result of such a reduction in the rural farm population, it is likely that the total farm income for the area would be decreased from $77,267,700 to $76,692,600 (columns *F.I.* and *F.′I.′*, Table II), a loss of $575,000, or a little less than one per cent. The weighted per capita income of the remaining farm population, however, would be raised about $38, or 15 per cent, as a result of an increase from $246 $\left(\text{i.e., } \frac{77{,}267{,}671}{314{,}078}\right)$ to $284 $\left(\text{i.e., } \frac{76{,}692{,}607}{270{,}124}\right)$ (columns *P*, *F.I.*, *P′*, and *F.′I.′*, Table II).

TYPE OF FARMING AREA II

Type of farming Area II (Tables III and IV) lies in southern and western Wisconsin, and also contains twenty-two counties. It is a region of dairy, livestock, and mixed farming. Farm land ranges in mean value from $28 to $46 an acre, with a mean for the area of $35 an acre. The weighted gross farm income per capita of the rural farm population is $199. It is somewhat poorer than Area I but more prosperous than Area III. The land-man ratio averages 28.8 acres per capita of the rural farm population, compared with 23 acres in Area I. The values $\eta^2_{02.13} = .33$; $\eta^2_{01.23} = .68$; and $H^2_{0.12(3)} = .71$ closely resemble the corresponding values found for Area I, and again indicate a strong relationship between per capita farm income and the land-man ratio. The estimates for

TABLE III

TYPE OF FARMING AREA II, TWENTY-TWO WISCONSIN COUNTIES: METHOD OF ANALYSIS*

X_2	X_1	20–25	26–31	32–37	38–41	Totals
25–39	n	7	7	1	2	17
	$\bar{X}_0$	181.1	203.9	196.0	245.0	198.9
	ΣX_0	1,268	1,427	196	490	3,381
	ΣX_0^2	231,394	292,999	38,416	120,212	683,021
	$(\Sigma X_0)^2/n$	229,689	290,904	38,416	120,050	672,421
40–49	n	1	2	2	0	5
	$\bar{X}_0$	181.0	224.0	245.5		224.0
	ΣX_0	181	448	491		1,120
	ΣX_0^2	32,761	101,152	120,292		254,205
	$(\Sigma X_0)^2/n$	32,761	100,352	120,541		250,880
Totals	n	8	9	3	2	22
	$\bar{X}_0$	181.1	208.3	229.0	245.0	205.0
	ΣX_0	1,449	1,875	687	490	4,501
	ΣX_0^2	264,155	394,151	158,708	120,212	937,226
	$(\Sigma X_0)^2/n_c$	262,450	390,652	157,323	120,050	930,448
	$\Sigma^r[(\Sigma X_0)^2/n_r]$					923,301
	$\Sigma^{rc}[(\Sigma X_0)^2/n_{rc}]$	262,450	391,256	158,957	120,050	932,718
	$(\Sigma X_0)^2/N$					920,864

* For meaning of symbols, see text.

TABLE IV

TYPE OF FARMING AREA II: OBSERVED DATA AND ESTIMATES[*]

No.	County	X_0	X_1	X_2	P	$F.I.$	I	X_0'	P'	$F.'I.' = X_0' \times P'$
1	Barron	189	21.6	36	21,916	4,012,329	146	258	11,446	2,953,068
2	Chippewa	183	25.8	33	18,131	3,326,164	135	236	11,785	2,781,260
3	Clark	187	20.5	37	24,533	4,588 151	149	260	12,512	3,253,120
4	Marathon	162	20.7	35	34,455	5,568,151	148	234	17,917	4,192,578
5	Pierce	211	25.5	37	14,432	2,827,877	136	265	9,236	2,447,540
6	Polk	167	23.5	34	18,953	3,173,630	141	229	11,182	2,560,678
7	Wood	169	22.3	37	13,725	2,319,452	144	213	7,686	1,637,118
8	Dunn	204	28.8	34	16,899	3,441,986	128	246	12,167	2,993,082
9	Eau Claire	181	28.5	31	10,352	1,878,082	129	224	7,350	1,646,400
10	La Crosse	236	30.8	32	8,877	2,091,438	123	271	6,835	1,852,285
11	Monroe	204	29.4	32	15,833	3,223,356	122	244	12,350	3,013,400
12	Pepin	183	28.8	31	4,677	853,836	128	225	3,367	757,575
13	St. Croix	214	28.5	30	15,018	3,208,562	129	257	10,663	2,740,391
14	Trempeleau	205	29.4	31	15,100	3,102,466	126	245	11,174	2,737,630
15	Crawford	196	35.2	28	9,323	1,824,384	112	214	8,204	1,755,656
16	Buffalo	254	40.3	28	10,134	2,572,740	100	254	10,134	2,572,740
17	Iowa	236	39.5	39	11,557	2,725,274	100	236	11,557	2,725,274
18	Vernon	181	25.5	42	19,193	3,476,233	125	229	14,395	3,296,455
19	Richland	204	29.5	40	12,146	2,479,452	113	229	10,567	2,419,843
20	Sauk	244	30.2	40	16,232	3,953,082	111	265	14,446	3,828,190
21	Grant	234	34.6	44	19,594	4,578,699	100	234	19,594	4,578,699
22	Lafayette	257	34.1	46	10,995	2,843,425	100	257	10,995	2,843,425
Totals					341,355	68,068,769			245,562	59,586,407
Means		199†	28.8	35			128†	243†		

[*] For meaning of symbols, see Table II.

† Weighted by the rural farm populations of the counties.

Area II are collected in Table IV. A land-man ratio of 40 acres was taken as the criterion for counties numbered 1 through 17; while for counties numbered 18 through 22 the criterion was 34 acres. In the first case the addition of each acre to the land-man ratio increased the per capita income by about $3.75, in the second case by about $5.62.

Area II shows a weighted index of farm population pressure of 128, compared with 114 for Area I. To reduce this index to 100 would require the removal of some 95,800 people from the farms of the twenty-two counties. Such a reduction in the farm population of the area would be accompanied by a loss in total gross farm income estimated at 12 or 13 per cent ($8,482,000). On the other hand, the weighted per capita income would be increased 22 per cent, from $199 to $243.

TYPE OF FARMING AREA III

The third and last type of farming area, covered by Tables V and VI, includes twenty-five counties. Nineteen of these counties fall in the northern part of the state, six in the central part. The area includes the northwestern sandy dairy-potato and dairy-hay regions, the central

TABLE V

TYPE OF FARMING AREA III, TWENTY-FIVE WISCONSIN COUNTIES: METHOD OF ANALYSIS*

X_2	X_1	14–19	20–25	26–31	32–37	38–47	Totals
10–24	n	3	5	3	1	2	14
	$\bar{X}_0$	83.0	118.4	144.3	170.0	153.5	125.1
	ΣX_0	249	592	433	170	307	1,751
	ΣX_0^2	20,981	71,110	64,769	28,900	48,220	233,980
	$(\Sigma X_0)^2/n$	20,667	70,098	62,496	28,900	47,125	219,100
25–39	n	3	7	1	0	0	11
	$\bar{X}_0$	116.0	136.1	165.0			133.3
	ΣX_0	348	953	165			1,466
	ΣX_0^2	43,322	135,863	27,225			206,410
	$(\Sigma X_0)^2/n$	40,368	129,744	27,225			195,459
Totals	n	6	12	4	1	2	25
	$\bar{X}_0$	99.5	128.8	149.5	170.0	153.5	128.7
	ΣX_0	597	1,545	598	170	307	3,217
	ΣX_0^2	64,303	206,973	91,994	28,900	48,220	440,390
	$(\Sigma X_0)^2/n_c$	59,400	199,068	89,400	28,900	47,125	423,893
	$\Sigma^r[(\Sigma X_0)^2/n_r]$						414,559
	$\Sigma^{rc}[(\Sigma X_0)^2/n_{rc}]$	61,035	199,837	89,721	28,900	47,125	426,618
	$(\Sigma X_0)^2/N$						413,964

* For meaning of symbols, see text.

TABLE VI

TYPE OF FARMING AREA III: OBSERVED DATA AND ESTIMATES*

No.	County	X_0	X_1	X_2	P	$F.I.$	I	X_0'	P'	$F.'I.' = X_0' \times P'$
1	Forest	84	19.6	22	3,116	263,219	156	125	1,371	171,375
2	Iron	70	14.7	22	2,995	210,411	167	119	988	117,572
3	Price	95	19.1	19	9,561	910,890	158	137	4,016	550,192
4	Ashland	123	21.6	23	5,086	624,109	152	161	2,441	393,001
5	Bayfield	126	23.2	22	8,338	1,053,014	148	161	4,336	698,096
6	Florence	126	24.3	20	1,664	209,315	146	160	899	143,840
7	Marinette	127	21.7	22	12,012	1,522,877	152	165	5,766	951,390
8	Sawyer	90	22.9	23	4,467	403,973	149	126	2,278	287,028
9	Burnett	134	27.4	21	7,574	1,014,041	139	163	4,620	753,060
10	Jackson	182	31.1	23	11,244	2,042,740	131	205	7,758	1,590,390
11	Washburn	117	28.7	21	5,860	683,151	136	143	3,750	536,250
12	Waushara	170	34.4	21	9,527	1,618,630	124	187	7,241	1,354,067
13	Adams	130	46.7	13	5,524	718,699	100	130	5,524	718,699
14	Marquette	177	43.6	18	5,781	1,020,479	100	177	5,781	1,020,479
15	Langlade	143	18.9	35	9,284	1,320,479	139	188	5,663	1,064,644
16	Taylor	133	19.7	28	12,519	1,666,849	136	175	8,012	1,402,100
17	Vilas	72	16.2	27	2,298	166,370	148	127	1,195	151,765
18	Douglas	111	21.6	25	7,870	873,836	130	146	5,509	804,314
19	Lincoln	134	23.3	29	8,195	1,101,781	125	163	6,146	1,001,798
20	Oconto	141	20.6	31	14,812	2,087,534	134	180	9,776	1,759,680
21	Oneida	104	25.2	27	3,936	407,945	119	126	3,188	401,688
22	Portage	127	25.8	28	16,256	2,063,904	117	146	13,492	1,969,832
23	Rusk	134	20.3	26	9,844	1,315,959	135	174	6,399	1,113,426
24	Waupaca	202	25.1	35	16,297	3,299,658	119	224	13,201	2,957,024
25	Juneau	165	31.4	26	9,690	1,597,329	100	165	9,690	1,597,329
Totals					203,750	28,197,192			139,040	23,509,039
Means		138†	25.1	24			132†	169†		

* For meaning of symbols, see Table II.

† Weighted by the rural farm populations of the counties.

sandy general farming region, and the central sandy-loam dairy-potato region. Farm land values are much lower than in either of the other two areas, ranging from $10 to $39 per acre, with a mean of $24 per acre. The weighted per capita farm income, only $138, is also the smallest among the three areas. The mean land-man ratio of 25 acres falls between the land-man ratios of Areas II and I. The values of the correlation ratios are somewhat less than those for the other two areas: $\eta^2_{01.23} = .49$; $\eta^2_{02.13} = .19$; $H^2_{0.12(3)} = .50$. However, there is again a substantial relationship between per capita farm income and land-man ratio. The criteria used were: among counties numbered from 1 through 14 in Table VI, a land-man ratio of 45 acres, with an increase in per capita income of $1.62 for each acre added to the land-man ratio; among counties numbered from 15 through 25, a land-man ratio of 31 acres, with an increase of $3.73 in per capita income for each acre added to the land man ratio.

The weighted index of farm population pressure for Area III is 132, which is the highest among the three areas. Some 64,700 persons would have to be withdrawn from the rural farm population in order that all counties in the area might attain the land-man ratios of the criterion counties. Such a population adjustment would result in an estimated reduction of about 17 per cent ($4,688,000) in the total gross farm income, but would raise the weighted mean farm income per capita 22 per cent, from $138 to $169.

CHANGES IN COUNTY FARM POPULATIONS, 1930–1940, IN RELATION TO 1931 INDEXES OF POPULATION PRESSURE

Between the censuses of 1930 and 1940 the rural-farm populations of the counties with indexes of population pressure greater than 100 in 1931 decreased about 3 per cent on the average in Area I and also in Area II; whereas the farm population of the criterion counties with indexes of 100 increased about 3 per cent in Area I and remained practically stationary in Area II. In type of farming Area III the farm population of the counties with indexes of population pressure above 100 increased nearly 5 per cent over the same period, while that of the criterion counties increased only about one per cent.

These figures suggest that in Area I adjustments in the size of the farm population have recently taken place that are tending to equalize population pressure on agricultural resources throughout the area at an average level above that of the criterion counties but below that of the counties with population indexes in excess of 100 in 1931. In type of farming Area II the slight recession of farm population in the criterion counties and the more marked decline in the other counties is also working toward equalization of pressure, but at a level below that of the criterion counties in 1931. By contrast, in Area III the inequalities of 1931 between the criterion counties and the more crowded counties have been increasing. This is further evidence that the major problem of population pressure on agricultural resources is to be found in the northern and central parts of the state, particularly in the so-called "Cut-Over" region.

SIZE OF FARM VS. SIZE OF FARM FAMILY IN RELATION TO INDEX OF FARM POPULATION PRESSURE

When the indexes of farm population pressure are linearly correlated with the percentage of young people under twenty years of age[11] in the rural-farm population in 1930 by counties in each type of farming area, and the three coefficients so obtained are averaged, we find $\bar{r}^2 = .25$. Similarly, between the indexes of population pressure and size of farm family[12] the mean $\bar{r}^2 = .14$; and between the indexes of population pressure and mean size of farm[13] the mean $\bar{r} = -.91$, or $\bar{r}^2 = .83$. If we represent the indexes of population pressure by X_4, the size of farm by X_5, and the size of farm family by X_6, and obtain coefficients of linear partial correlation, we find $\bar{r}_{45.6} = -.86$, or $\bar{r}^2_{45.6} = .74$; and $\bar{r}_{46.5} = .49$, or $\bar{r}^2_{46.5} = .24$. Finally, the simple linear correlation between size of farm and size of farm family is mean $\bar{r} = -.20$, or $\bar{r}^2 = .04$. The interpretation of these coefficients has been given above under "General Method and Results."

[11] 1930 Federal Census, *Population*, Vol. III, Pt. 2, p. 1307.

[12] Ibid., *Agriculture*, Vol. II, Pt. 1, p. 736.

[13] Ibid., *Population*, Vol. VI, *Families*, p. 1468.

NOTE: It will be noticed that all of the data used in this study were taken from the 1930 Federal Census, except the gross farm income by counties, which was furnished by the Wisconsin Department of Agriculture and Markets. Unfortunately, reliable estimates of farm income by counties are rarely available in any state, including Wisconsin. A fairly good substitute, however, may now be obtained from the Federal Census. This is the value of all farm products sold, traded, or used by farm households (*Agriculture*, Third Series, County Table XVII, 1940). The squared coefficient of correlation between these values of farm products and the gross farm incomes (X_3) used in this study, both on a per capita basis, is $r^2 = .81$, as already mentioned.

A study dependent on county data drawn solely from the Federal Census and publications of a state department of agriculture, of course, should not be used alone to recommend or guide any actual adjustments of the farm population to the land resources of a county or state. When such adjustments are made, as in the case of the resettlement projects of the Federal Government in northern Wisconsin, such studies as this should be supplemented by intensive economic, farm management, and sociological studies conducted in the field.

THE PRODUCTION FUNCTION FOR CANADIAN MANUFACTURES[*]

BY PATRICIA DALY
University of Chicago
AND PAUL H. DOUGLAS
Captain, U.S.M.C.R., and *University of Chicago*

READERS of this JOURNAL may perhaps be somewhat acquainted with a series of articles by one of the present authors and various associates which on the basis of (1) time-series and (2) cross-section studies of labor, capital, and product have sought to discover by use of the production formula, $P=bL^kC^j$, the effects upon the quantity of manufacturing product caused by changes in the quantities of labor and capital. Thus far the economies studied have been the manufacturing industries of the United States[1] and Australia.[2] The present study gives empirical values to the parameters of the above function for the Canadian manufacturing industries. Four years have been chosen for the cross-section analysis,[3] namely, 1923, 1927, 1935, 1937, care being taken to avoid depression years and to select years which were relatively "normal."

In each of these given years, each manufacturing industry was treated as an observation with its respective quantities of Labor (L), Capital (C), and Product (P). The definitions given to these terms were as follows:

1. Labor (L) = total number employed, i.e., wage-earners and salaried workers.
2. Capital (C) = total capital, including both (a) fixed capital in plant, tools, machinery, etc., and (b) working capital including materials, goods in process, etc.

---

[*] The authors wish to thank their friends and colleagues, Miss Grace T. Gunn and Mr. H. Gregg Lewis for numerous helpful suggestions. We also wish to acknowledge aid given, in computing by E. Olson, B. Nimer, and W. L. Slayton, and in drawing of the charts by Mr. Y. K. Wong.

[1] Paul H. Douglas, *The Theory of Wages*, pp. 113–167; M. Bronfenbrenner and Paul H. Douglas, "Cross-Section Studies in the Cobb-Douglas Function," *Journal of Political Economy*, XLVII (December, 1939), pp. 761–785; Grace T. Gunn and Paul H. Douglas, "The Production Function for American Manufacturing in 1919," *American Economic Review*, XXXI (March, 1941), pp. 67–80; and "The Production Function for American Manufacturing for 1914," *Journal of Political Economy*, L (August, 1942), pp. 595–602; Patricia Daly, Ernest Olson, and Paul H. Douglas, "The Production Function for Manufacturing in the United States, 1904," *Journal of Political Economy* (forthcoming).

[2] Grace T. Gunn and Paul H. Douglas, "Further Measurements of Marginal Productivity," *Quarterly Journal of Economics*, LV (May, 1940), pp. 399–428; and "The Production Function for Australian Manufacturing," *ibid.*, LVI (November, 1941), pp. 108–129; Paul H. Douglas, *The Theory of Wages*, pp. 167–172; Paul H. Douglas and Marjorie L. Handsaker, "The Theory of Production as Derived from Australian Data," *Quarterly Journal of Economics*, LII (November, 1937, and March, 1938) pp. 1–36, 215–54.

[3] For the sources, see the Canada Year Book for 1925, 1930, 1938, and 1940.

3. Product (P) = gross sales value of product minus cost of materials used, fuel, electricity, etc.

The equation $P = bL^kC^j$ can be rewritten as a linear one by using the logarithmic values of the variables. The equation then becomes

$$\log P = \log b + k \log L + j \log C.$$

Fitting this equation by the method of least squares, the values in Table I are obtained.

It will be noticed:

1. That the sum of the exponents for labor and capital (i.e., $k+j$) closely approximate unity or 1.0. This is what has been found to be approximately true in virtually all of the other studies and strengthens the conclusion that the production function is homogeneous and of the first degree. This corroborates the deductive conclusion drawn that under competition production is carried to the point of lowest cost on the average total unit cost curve which is the point of constant costs.

2. That the exponents of both labor (k) and capital (j) are relatively steady. The span in values of k was from .43 to .50 which was less than the range of one standard error from the intermediate values. The span in values of j was from .48 to .58.

3. The values of k are somewhat less and the values of j somewhat more than those for the United States found for four particular years of the period 1904–1919 which were as follows:

Year	k	j
1904	.65	.31
1909	.74	.32
1914	.61	.36
1919	.76	.25

This is not surprising. We have never claimed that the exponents of the factors should be the same for all economies and periods. Indeed in *The Theory of Wages*, it was explicitly stated that one might expect such variations.[4] In the United States where capital per worker is appreciably greater we might expect capital to have a lower and labor a higher exponent than in Canada.

It is also interesting to note that the exponents of labor and capital for New Zealand as shown by a time study by Dr. Max Brown seem to approximate those of Canada.[5]

---

[4] See, Paul H. Douglas, *The Theory of Wages*, p. 203.

[5] The results of this study are given by Colin Clark in his book, *Conditions of Economic Progress*, p. 384. The material is contained in an unpublished thesis at Cambridge University, which through the courtesy of Dr. Brown we have had the privilege of reading.

We would conclude that the elasticity of the marginal productivity curve for labor in Canadian manufacturing for the four years studied was probably slightly under 2.0 and that for capital slightly over 2.0 (i.e., under competitive conditions the elasticity of this curve is equal to the reciprocal of the exponent of the other factor).

A test of the accuracy of the production function is the following: (1) compute for each industry from its actual quantities of Labor (*L*) and Capital (*C*) what the Product should be according to the formula. We shall call this *P'*; then (2) compare the actual product (*P*) with this computed product (*P'*) and see how great are the relative differences of the former from the latter and whether they can be explained by random errors of measurement or by other random causes operating upon the value product of specific industries.

TABLE I

VALUES OF EXPONENTS OF LABOR AND CAPITAL IN THE PRODUCTION FUNCTION FOR CANADIAN MANUFACTURES*

Year	No. of Industries	k	σ_k	j	σ_j	b	$k+j$
1923	167	.48	.038	.48	.035	48.5	.96
1927	163	.46	.039	.52	.036	33.0	.98
1935	165	.50	.044	.52	.040	22.2	1.02
1937	164	.43	.041	.58	.036	15.4	1.01
Average for four years		.468		.525			.992

* If we use fixed capital only for *C*, the exponents are changed to the following:

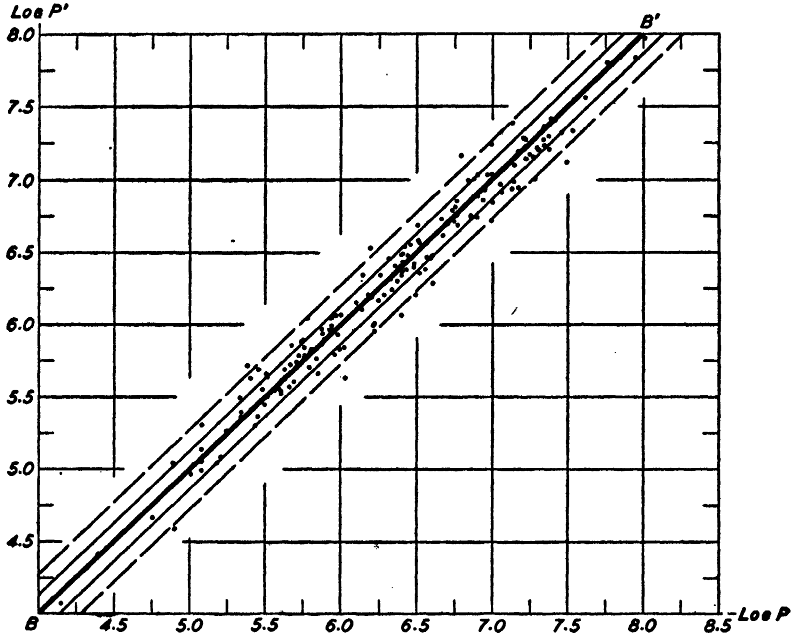
Year	k	j	b	$k+j$
1923	.54	.44	82.0	.98
1927	.49	.49	58.9	.98
1935	.53	.49	33.5	1.02
1937	.48	.53	32.9	1.01
Average	.510	.488		.998

This has been done for the years 1923, 1927, and 1935; the results being shown graphically in Charts I to III inclusive. The logs of the actual product or *P* are shown on the horizontal axis or abscissa; the logs of the computed product or *P'* upon the vertical axis or ordinate. Were there to be perfect correspondence between the two, each observation would fall on the line *BB'*. The differences between the actual product (*P*) and those which we would expect under the formula (*P'*) are measured by the horizontal deviations of the observations from the line *BB'*. To measure these more accurately two sets of bands have been laid out on the charts. The first is at a distance of one standard error of estimate to the right and to the left of the line *BB'* while the second set of bands is at a distance of from one to two standard errors of estimate from the line of expected relationship.

It will be noticed that the observations cluster closely around the line BB' and that the scatter is not great. This is further evidence that the production formula furnishes a good description of the product in these manufacturing industries. In Table II there is shown the percentage of observations deviating from the line of perfect correspondence by less than one and two, and by more than two standard errors of estimate of P'. Chart IV is a graph of the results for 1937.

CHART I

THEORETICAL AND OBSERVED VALUES OF LOG P FOR 1923



We should not expect that there would be a precise agreement in each industry between P and P'. The modern theory of causation does not require any such one to one relationship. It is enough that given causal forces express themselves in "norms" of results or tendencies accompanied by a more or less symmetrical range of variations resulting from errors of sampling, errors of measurement, and sheer chance. That is one of the implications of the theory of probability. If the actual results do approximate the norms fairly closely and the variations are rather symmetrically distributed, then credence in the formula as both descriptive and predictive is greatly strengthened.

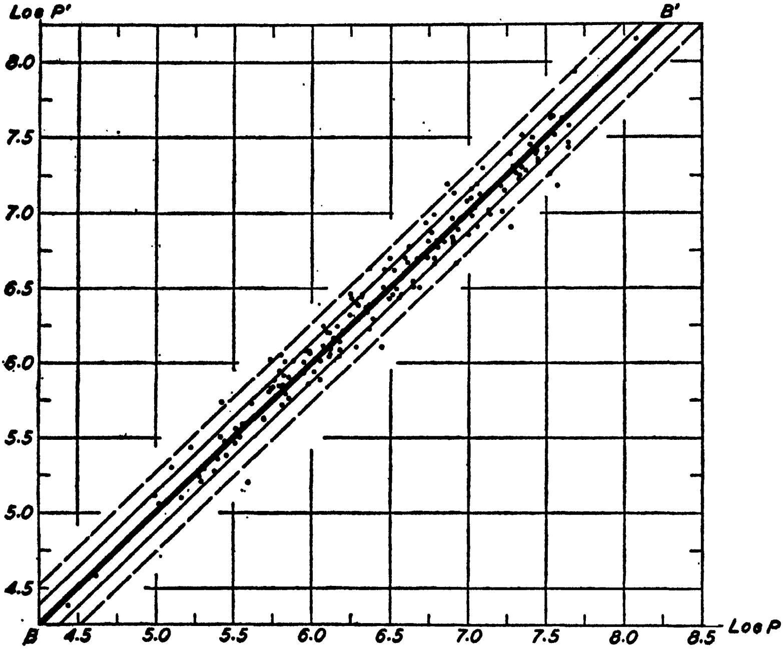
In the present case we would not expect deviations due to sampling errors because of the fact that the quantities of L, C, and P are all derived from an all inclusive census and not from any sample. Some differences between P and P' would however inevitably result from:

(1) Errors of measurement.

(2) Shifts over time in the demand schedules for the end products of the industries causing some to move to the right and others to the left and hence affecting the money values of the products. Here it should be noted that under the quantity theory of money, if the supply of money and bank credit is constant, the average elasticity of demand is equal to unity. So is its reciprocal, the average flexibility of price. The variations in P caused by shifts of the demand curves will under these conditions tend to be mutually compensatory and hence will operate in a non-biased manner. The deviations from the norm caused by such shiftings will have, therefore, some tendency towards symmetry.

CHART II

THEORETICAL AND OBSERVED VALUES OF LOG P FOR 1927



(3) Differences in the technical effectiveness of labor and capital in specific industries. We have never maintained that labor and capital must have the same exponents in each and every industry, but have

TABLE II

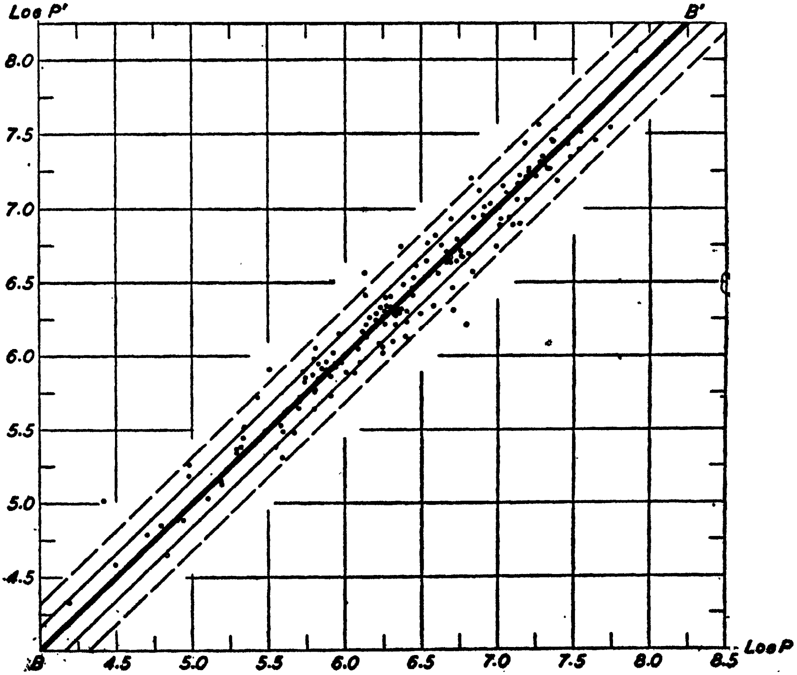
DEVIATIONS OF ACTUAL PRODUCT (P) IN CANADIAN MANUFACTURING FROM PRODUCT (P') COMPUTED UNDER THE FORMULA FOR THE YEARS 1923, 1927, 1935, AND 1937

Year	Percentage of observations of given deviations from P'		
	Within $\pm 1S$	Within $\pm 2S$	Outside $\pm 2S$
	Per cent	Per cent	Per cent
1923	72.46	92.81	7.19
1927	74.23	95.09	4.91
1935	72.17	95.76	4.24
1937	76.83	93.90	6.10

instead always pointed out these may well differ. The empirical exponents which we find are instead merely norms or averages of these varying exponents. It would be expected, therefore, that the actual products from industry to industry would not jibe precisely with the computed products based on the use of these "average" exponents of labor and capital. But, since the exponents for manufacturing industry as a whole are the norms or averages for the specific industries, the

CHART III

THEORETICAL AND OBSERVED VALUES OF LOG P FOR 1935

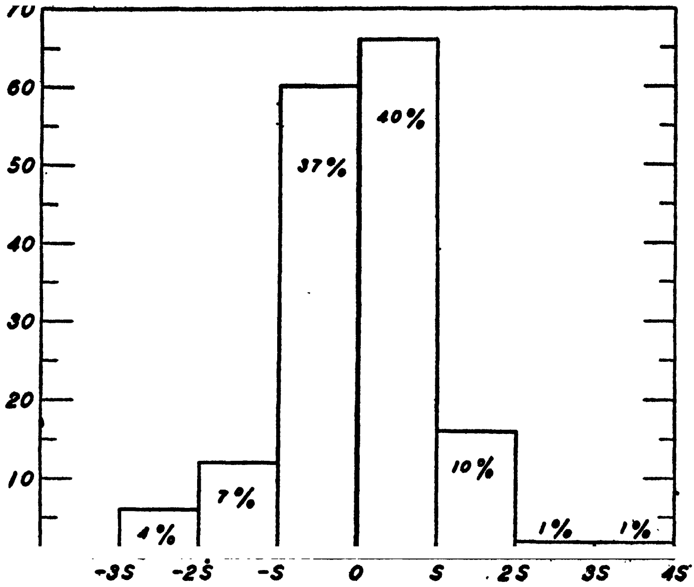


variations of these actual exponents of labor and capital in the individual industries would again be more or less symmetrically distributed and would not be biased.

(4) A fourth factor in causing differences between actual and computed P would be forces affecting production other than the quantities of labor and capital.

CHART IV

FREQUENCY DISTRIBUTION OF DEVIATIONS OF P FROM P' FOR 1937



(5) Finally, sheer chance itself may also be expected to play a part in causing individual deviations from the norms.

But, and this is crucially important, to the degree that these other causes operate at random and not in a biased manner, then the values of the exponents of labor and capital in the production function are the main determinants of the end result, namely, the value of the product. Under these conditions, we would expect that in two-thirds of the cases, the actual product would not deviate from the theoretical product by more than one standard error and that in 95 per cent of the cases, these deviations would not be more than two standard errors.

In practice we find that about 74 per cent of the observations fall within one standard error of the theoretical production values, and that in approximately 94 per cent of the cases, the deviation is less than two standard errors. The relative differences between the actual and computed values of P are therefore at least not more, and in part actually less than would be expected under the theory of normal probability. Credence in the developed production function as a description of the relative effects of labor and capital upon product is thereby strengthened. The variations from it are normal.

A study of the relatively small number of industries where the deviations of actual from theoretical products were more than two standard errors of estimate shows that in the case of charcoal and coke, the actual product fell short of the theoretical by this amount in no less than three out of the four years; so too did clay products, linen goods, agricultural implements, and petroleum products in two out of the four years. On the other hand the actual values of abrasive products exceeded the theoretical values by more than two standard errors in three out of the four years. Also the actual values of cigars and cigarettes exceeded the theoretical values by more than two standard errors in two out of the four years. In view of the high degree of monopoly prevailing in this industry this is not surprising. Similarly in the case of breakfast foods, the money values of the actual products also exceeded the theoretical values by more than two standard errors in two of the years. In view of the importance of trade-marked brands creating a state of highly imperfect competition this again is what should be expected. Another industry where in two of the years the actual products were more than two standard errors above the theoretical values was the liquor industry.[6]

Some of the industries for which in one of the four years the actual value of product differed from the theoretical by more than two standard errors were toys, chopping mills, coal tar, fly paper, signs, paper patterns, bicycles and motorcycles, fur dressing and dyeing, maple syrup and sugar, fertilizers, musical instruments, fountain pens, nonferrous metals, sand lime brick, and wood distillation. On the whole a study of the deviations gives added proof of the general reliability of the formula and of its values for the four years studied.

A further test of the function can be made by comparing the actual share of the value product received by labor with that which we would expect to be received with the production formula. Under conditions of perfect competition one would expect labor to receive as its proportionate share of the total product the proportion represented by its own exponent k or more properly by the ratio $\frac{k}{k+j}$; $\frac{k}{k+j}$ being a more proper representation of labor's share of product under conditions of increasing or decreasing returns where $k+j$ is greater than or less than one.

---

[6] Distilled liquors in one year and malt liquors in another.

In Table III we make this comparison, the share of labor being represented by W/P or the ratio of total wages and salaries to the net value of the product. Taking the four years as a whole there was an almost precise correspondence between the share which labor actually received and that which we would expect it to receive. In the first two years, this agreement was very close, while in the last two years there were compensating differences. In 1935, the share which labor received was from 9 to 10 per cent less than one would expect from the formula while in 1937 it was 9 per cent more.

TABLE III

Year	W/P	k	$\frac{k}{k+j}$
1923	.50	.48	.50
1927	.48	.46	.47
1935	.40	.50	.49
1937	.52	.43	.43
Average	.475	.468	.472

Taken in the large, the results of our Canadian studies seem to strengthen the case for the use of the production function.

THE COMPOSITION OF INCOME AND OWNERSHIP OF CAPITAL BY INCOME CLASSES IN THE UNITED STATES IN 1936

By Rufus S. Tucker
General Motors Corporation

Much current controversy revolves around the point of the distribution of income and wealth, and the alleged necessity of changing it, by taxation or other means. For that reason any statistical information concerning their distribution, if it is fairly close to accuracy, should be helpful as a substitute for prejudices and assumptions. A good deal of information relating to the source of income in the various income classes can be gleaned from the recently published special studies of income tax returns for 1936, filed in accordance with the income tax laws of the United States and of the states of Wisconsin and Delaware. The chief defects of the Federal statistics—the classification by statutory net income and the large proportion of persons not required to file returns—are remedied by the state statistics, which are classified by total income less capital gains and losses, and which cover a larger proportion of the population. Delaware, in fact, requires returns from every adult, irrespective of the size of his income.

Wisconsin and Delaware are perhaps more prosperous than the average state. Since, however, one is mainly agricultural and the other mainly industrial, they are not likely to deviate in the same direction from the national average in the matter of the composition of income within a given income class. For example, it is not likely that both would show a smaller proportion of the income of the $1,500–$2,000 class derived from labor, or a larger proportion from property, than the average of all the other states, although they might both show that a larger or smaller proportion of all persons fell within the $1,500–$2,000 class than would be the case in other states. For families with incomes under $2,500 and single individuals with incomes under $1,000, conclusions derived from these state statistics are more likely to be correct than those derived from the Federal statistics, since the coverage is better and there is less bias on account of filing requirements and other legal technicalities.

From these statistics it is possible to calculate the proportion of persons in each income class who received various kinds of income, and the proportion of their income received from each of the sources. Applying these proportions to the estimated numbers of persons and of income in each class we can build up aggregates showing the distribution of each kind of income, hence the ownership of various kinds of capital. These figures can be worked out separately for families and single individuals; in some cases separately for single men and women.

The family figures can be supplemented by the results of the Consumer Purchases Study, which were published in a large number of bulletins by the Bureau of Labor Statistics and the Bureau of Household Economics, and summarized to some extent in the National Resources Committee volumes called *Consumer Expenditures* and *Family Expenditures*. Some of these data refer only to non-relief families and may therefore overstate the importance of some forms of property income in the lower income classes. But if corrected, as they can be in many cases, for relief families, they may tend to understate property incomes, because they were derived to a large extent from families that either were poorer than the average or else understated their income from property.[1]

The CPS data do not state separately profits from individually owned enterprises such as farms, although they make a general distinction between farmers (some of whom earned some wages) and wage earners (some of whom made some profits). Delaware is not so much of a farming state as most, and in addition farmers were believed to be most likely to neglect to file returns. Also in all states business and professional men (including farmers) are generally supposed to be most likely to understate their incomes, as there is less check on net profits than on other forms of income. Consequently the ratios derived from tax returns probably underestimate the importance of profits in all income classes except those large enough to cause the Treasury to make a special audit. No correction on this account has been possible. Wages, partnership income, rent and interest were reported at the source if they exceeded the amount of the individual or family exemption, and are therefore fully represented in the Federal individual tax returns above say $3,500 for families and $2,000 for single individuals. In Delaware they were reported at the source if in excess of $1,000, irrespective of the recipient's family status.

Dividends amounting to $100 or more in a year were reported at the source under the Federal law and were therefore fully reported by persons filing returns, but not necessarily reported at all by families under

---

[1] For evidence on this point see my articles in the *Review of Economic Statistics*, November 1940 (Vol. XXII, pp. 165–182) and February 1942 (Vol. XXIV, pp. 9–21). Further evidence of understatement (or omission) is found in the fact that the number reporting savings bank accounts far exceeded the number reporting the receipt of either interest or dividends, in every income class. Also, in the income classes above $2,500, where the coverage of the tax returns is complete, the proportion of families reporting dividends and the proportion reporting interest on tax returns, each exceeded the combined proportion of families reporting either dividends or interest on the CPS returns. Likewise the proportion of total income derived from dividends or interest reported on the tax returns each exceeded the proportion of dividends and interest together reported in the CPS, for every income class.

$2,500 or single individuals under $1,000. Consequently a large amount of dividends is omitted, since it was received by persons in the lower income brackets who filed no Federal returns.

In general the tax ratios may slightly exaggerate the importance of property income and minimize the importance of individual profits in all classes between $2,500 and $20,000. Below $1,000 one might suspect that they exaggerated the importance of property incomes, but later on in this article evidence is given to show that that was not the case with respect to the Delaware ratios which have been used for family cash income in classes below $2,500. To these ratios the NRC estimates of imputed rental from owned homes and home-produced food and fuel have been added, since these items were not included in the Delaware returns.

Of course the ratios from the various tax tabulations do not agree exactly; in some cases there are wide variations. But these variations can usually be explained by differences in the legal definition of the kind of income under consideration, or of the income taken as the basis for tabulation. Or else they are caused by differences in the extent of coverage, or by the inclusion of separately reporting husbands and wives in their individual income classes and not in the class where the family belongs. Fortunately it has been possible in every calculation to eliminate the separately reporting husbands and wives, and in most cases to put the combined family in its proper class. If, after this adjustment was made there were two ratios that seemed equally reasonable in the light of the definitions and coverage, that one was taken which showed the lower proportion of property income below $2,500, and the higher proportion above; and the opposite choice was made for labor income.

Besides the tables showing the ratios of all persons receiving income from the different sources, and the total amount received, there are official tables of "Patterns" of income, showing the number of persons wholly relying on income from each source. These "Patterns" are in many respects unsatisfactory; they all lump together families, single individuals and separately reporting husbands and wives. The United States "Patterns" includes estates and trusts. The Wisconsin "Patterns" has excessively wide limits to its income classes. Both the United States and the Wisconsin "Patterns" are tabulated according to a different definition of income from that used in the other tables. The United States "Patterns" tabulates the sources in such a way as to conceal the nature of the fourth source of income and to conceal the number of those having five or more sources. Wisconsin conceals the nature of the third source of income and the number of those having four or more sources. None of the "Patterns" include any imputed income, except an infinitesimal amount of imputed profits in Wisconsin.

Consequently the number of incomes derived wholly from labor, or profits, or property, or profits and property together, is exaggerated by these "Patterns." A considerable amount of correction is possible, by using other tax tables and those in the Consumer Purchases Study, but the figures used in this study probably still overstate the number of single-source incomes, especially single-source property incomes.

The estimate of income distribution to which these ratios were applied was based on the one I published in this JOURNAL last December, but adjusted to the calendar year 1936. If the National Resources Committee's distribution were used as a base it would increase both the number of property incomes and the number of labor incomes in the classes below $1,000 and above $7,500, and would give a larger total amount of property income and a larger number of persons receiving only property income. As I stated in that article, I believe that both the NRC figure and my own, for the under $500 class, were too large.

The definition of income used for classification is the same as that used by the National Resources Committee in its studies, with the inclusion of imputed income for families as estimated by the NRC, but no allowance for imputed income of single persons. There is, therefore, a slight exaggeration of the importance of labor income in the lower classes. Capital gains are not included in income in this analysis.

We may now proceed to a consideration of the detailed tables, and the more important conclusions to be drawn from them.

LABOR (WAGES AND SALARIES)

In the first place the commonest source of income is labor. That is not a startling discovery, but the distribution of labor income by classes is perhaps a little unexpected. The greatest proportion of labor income is found, not in the lowest class (under $500) but in the class $1,000–$1,500, where 90 per cent of the consumer units report it. The proportion declines with increasing income until 71.7 per cent is reached in the $10,000–$15,000 class, but above $15,000 over 75 per cent report labor incomes.

Taking the proportion of labor income to all income: that also is highest in the $1,000–$1,500 class and is higher in every class up to $3,500 than it is in the lowest class (under $500), where it is only 65.7 per cent. Even above $20,000 the proportion of labor income was 22 per cent. Except in the $1,000–$2,000 class, families were more likely than single persons to have labor incomes, and in every class single men were more likely than single women to have labor incomes.

Considering now the proportion of consumer units wholly dependent on labor income, here again the lowest class does not show the highest

TABLE I

PER CENT RECEIVING ANY INCOME FROM SPECIFIED SOURCES, UNITED STATES, 1936
(Consumer units)

Income class	Labor	Business (cash)	Partnership	Imputed profits	Any profits (min.)	Any property or profits (min.)
0– 500	78.2	7.6	.4	15.6	17.8	34.3
500– 1,000	88.2	8.0	.4	18.0	21.4	24.4
1,000– 1,500	90.0	8.4	1.7	16.0	22.3	24.6
1,500– 2,000	89.0	13.7	2.1	14.1	23.8	36.3
2,000– 2,500	86.7	19.0	2.8	11.5	23.2	54.2
2,500– 3,000	83.6	19.2	4.2	12.1	25.7	56.3
3,000– 3,500	81.4	21.4	5.7	11.7	28.0	64.3
3,500– 4,000	80.3	23.7	6.8	11.9	28.9	75.5
4,000– 4,500	79.1	26.7	7.8	10.8	33.6	78.1
4,500– 5,000	77.9	28.6	7.6	10.6	35.9	81.2
5,000–10,000	75.0	34.5	10.5	10.7	46.2	89.2
10,000–15,000	71.7	40.7	13.9	7.0	51.6	96.5
15,000–20,000	75.7	44.6	14.7	7.0	56.5	96.9
20,000 & over	75.3	50.8	18.3	6.9	58.5	99.8

Income class	Fiduciary & other	Interest	Rent (cash)	Imputed rent	Dividends	Any property (min.)
0– 500	.4	17.6	6.7	17.6	7.4	26.8
500– 1,000	.4	10.0	5.0	19.5	5.3	19.5
1,000– 1,500	.5	13.1	5.0	24.6	7.1	24.6
1,500– 2,000	.7	19.4	7.0	31.0	10.6	31.0
2,000– 2,500	1.1	22.2	10.8	38.5	18.7	43.3
2,500– 3,000	1.8	27.2	13.9	47.6	20.4	52.0
3,000– 3,500	2.1	33.3	17.7	47.7	25.0	52.9
3,500– 4,000	2.6	34.2	21.9	46.8	31.9	61.4
4,000– 4,500	3.6	35.2	23.9	55.0	38.0	66.1
4,500– 5,000	4.2	36.5	24.5	55.5	42.1	63.6
5,000–10,000	7.5	44.1	27.4	54.4	47.3	74.9
10,000–15,000	11.1	56.3	28.5	48.9	71.0	83.6
15,000–20,000	13.5	62.4	28.1	48.0	78.9	91.4
20,000 & over	29.3	74.1	32.7	47.2	87,3	95.2

proportion, but is surpassed by the classes between $500 and $1,500. The highest proportion is in the class $500–$1,000, if cash income alone is considered, and in the class $1,000–$1,500 if allowance is made for imputed income. In most classes families appear to be more likely than single individuals to depend wholly on labor income. Perhaps that is because young unmarried persons can save more easily than married ones, and aged persons who have acquired property during their prime are likely to be widowed. In every class single men are much more likely than single women to be dependent on labor only.

If the ratios presented in Table I are applied to the distribution of income presented in my article in the December JOURNAL, the number of consumer units receiving some labor income comes out as 35,433,000. This represents a considerably larger number of workers since working wives and children are not counted. The total of gainfully employed reported by the 1930 Census was 36,470,000 after deducting farm owners and operators and the unpaid members of their families. Plainly the number of hired workers has not been underestimated in this study.

Similarly the total of wages and salaries obtained by using the ratios of income shown in Table III is $45,387,000,000, while the amount calculated for 1936 by the Department of Commerce is $42,125,000,000.

TABLE II

PER CENT RECEIVING ANY INCOME FROM SPECIFIED SOURCES, UNITED STATES, 1936

Income class	Families*			Single persons†		
	Labor	Profits (min.)	Property (min.)	Labor	Profits (min.)	Property (min.)
0– 500	83.5	30.9	29.5	75.0	3.4	26.8
500– 1,000	89.0	32.4	29.1	87.0	4.9	13.4
1,000– 1,500	90.0	29.0	33.4	92.6	3.8	18.4
1,500– 2,000	88.3	27.2	36.9	90.0	5.7	29.1
2,000– 2,500	86.7	25.4	44.6	84.7	9.4	43.3
2,500– 3,000	82.0	27.1	52.0	80.1	12.8	53.3
3,000– 3,500	78.8	29.3	52.9	74.7	16.2	64.5
3,500– 4,000	76.3	30.5	52.9	63.7	17.6	64.5
4,000– 4,500	75.0	35.6	62.8	61.1	18.8	75.3
4,500– 5,000	77.4	38.2	62.8	57.8	19.1	75.3
5,000–10,000	73.8	50.2	69.1	54.8	20.2	82.3
10,000–15,000	73.6	51.6	82.0	55.1	18.5	92.6
15,000–20,000	77.4	56.5	90.5	53.3	18.4	86.9
20,000 & over	81.3	58.5	96.0	52.7	13.2	94.2

Income class	Single men†			Single women†		
	Labor	Profits (min.)	Property (min.)	Labor	Profits (min.)	Property (min.)
0– 500	87.8	4.2	13.8	64.7	2.7	34.4
500– 1,000	91.6	5.6	8.6	82.2	4.1	18.5
1,000– 1,500	93.2	4.5	12.6	87.6	3.1	25.5
1,500– 2,000	92.8	6.0	20.7	80.6	5.0	42.5
2,000– 2,500	87.3	9.8	29.4	74.3	8.4	60.3
2,500– 3,000	84.0	13.9	41.7	65.1	10.8	67.4
3,000– 3,500	78.1	18.3	56.9	50.0	12.5	73.0
3,500– 4,000	78.1	20.0	56.9	50.0	14.0	73.0
4,000– 4,500	82.0	22.5	66.0	31.9	13.5	85.1
4,500– 5,000	82.0	23.1	66.0	31.9	13.5	85.1
5,000–10,000	69.5	25.5	80.0	20.9	13.7	84.9
10,000–15,000	55.1	24.4	88.9	5.3	12.3	96.3
15,000–20,000	80.0	25.6	80.0	27.3	11.5	92.1
20,000 & over	53.8	20.5	92.3	17.0	8.0	95.7

* Including imputed rent and profits.

† Excluding imputed income.

TABLE III

PER CENT OF INCOME RECEIVED FROM SPECIFIED SOURCES, UNITED STATES, 1936
(Consumer units)

Economic income class	Labor	Business profits (cash)	Partnership profits	Profits (imputed)	All profits	All profits or property
0- 500	65.7	4.3	.3	10.4	15.0	34.3
500- 1,000	75.6	5.7	.3	8.2	14.2	24.4
1,000- 1,500	79.2	6.4	.4	5.4	12.2	20.8
1,500- 2,000	78.0	8.4	.8	3.9	13.1	22.0
2,000- 2,500	75.0	10.0	1.5	2.7	14.2	25.0
2,500- 3,000	72.0	10.6	2.2	2.6	15.4	28.0
3,000- 3,500	68.7	12.3	2.8	2.1	17.2	31.3
3,500- 4,000	64.4	13.9	3.5	2.0	19.4	35.6
4,000- 4,500	60.4	15.1	4.3	1.4	20.8	39.6
4,500- 5,000	58.2	15.1	4.7	1.3	21.1	41.8
5,000-10,000	47.5	14.9	6.2	.9	22.0	52.5
10,000-15,000	39.4	12.7	7.8	.4	20.9	60.6
15,000-20,000	34.3	11.1	8.4	.3	19.8	65.7
20,000 & over	22.0	5.6	9.5	.1	15.2	78.0

Economic income class	Fiduciary & other property	Interest	Rent (cash)	Rent (imputed)	Dividends	All property
0- 500	.4	6.1	4.6	5.6	2.6	19.3
500- 1,000	.4	2.2	2.3	4.0	1.3	10.2
1,000- 1,500	.5	1.4	1.4	4.2	1.1	8.6
1,500- 2,000	.7	1.3	1.5	4.2	1.2	8.9
2,000- 2,500	1.0	1.8	2.0	4.2	1.8	10.8
2,500- 3,000	1.8	1.8	2.2	4.2	2.6	12.6
3,000- 3,500	1.8	2.0	2.7	4.2	3.4	14.1
3,500- 4,000	2.6	2.5	2.7	4.1	4.3	16.0
4,000- 4,500	3.6	2.9	2.6	4.0	5.7	18.8
4,500- 5,000	4.2	3.2	2.6	4.0	6.7	20.7
5,000-10,000	5.4	9.2	2.5	3.5	9.8	30.5
10,000-15,000	8.0	11.3	1.5	3.1	15.8	39.7
15,000-20,000	9.1	12.3	1.5	2.9	20.1	45.9
20,000 & over	11.4	11.7	1.8	2.5	35.5	62.8

The total receiving no cash income except wages or salaries was about 25,000,000 according to these calculations; if an allowance were made for rent of owned homes and imputed profits received by families the total might be as low as 11,000,000. For comparison the total number of consumer units in the United States in 1936 is estimated to have been 41,234,000. But, as will be shown later, these persons with no income but wages and salaries cannot be regarded as a propertyless proletariat. Many of them had high incomes, and most of the families with low incomes owned property, such as automobiles, radios and other durable consumers' goods, and insurance policies, and bank accounts. Moreover many of the single persons owned their homes although it has been impossible to include their imputed rental in these tables.

PROFITS, OR ENTREPRENEURIAL INCOME

Business and partnership profits include the income of farmers, professional people, and unincorporated merchants and manufacturers. They stand between wages and salaries, which represent payment for labor, and the various kinds of property income which are payments for the use of capital. They are a joint product of labor and capital; the persons who receive them are both laborers and capitalists. Individual

TABLE IV

PER CENT OF INCOME RECEIVED FROM SPECIFIED SOURCES, UNITED STATES, 1936

Income class	Families*			Single persons†		
	Labor	Profits	Property	Labor	Profits	Property
0– 500	61.6	23.1	15.3	75.0	3.4	21.6
500– 1,000	71.3	19.3	9.4	85.3	4.3	10.4
1,000– 1,500	76.8	15.0	8.2	90.4	3.1	6.5
1,500– 2,000	77.2	14.6	8.2	84.5	4.5	11.0
2,000– 2,500	75.0	15.4	9.6	73.8	7.0	19.4
2,500– 3,000	73.0	15.8	11.2	62.3	9.5	28.2
3,000– 3,500	70.0	17.7	12.3	57.3	11.9	35.7
3,500– 4,000	66.0	20.1	13.9	47.1	12.9	41.1
4,000– 4,500	62.4	21.6	16.0	40.5	13.6	45.9
4,500– 5,000	60.6	22.0	17.4	38.0	14.0	48.0
5,000–10,000	50.2	23.2	26.6	32.7	13.3	60.2
10,000–15,000	42.6	22.4	35.0	27.0	11.3	70.6
15,000–20,000	37.8	21.3	40.9	23.7	10.9	74.2
20,000 & over	24.6	16.6	58.8	19.5	8.0	81.7

Income class	Single men†			Single women†		
	Labor	Profits	Property	Labor	Profits	Property
0– 500	86.4	4.2	9.4	64.7	2.7	32.3
500– 1,000	90.0	5.3	4.7	80.1	3.4	16.4
1,000– 1,500	92.5	3.7	3.8	87.6	1.8	10.6
1,500– 2,000	89.2	5.0	5.8	74.7	3.0	22.3
2,000– 2,500	82.0	7.9	10.1	55.4	4.9	39.7
2,500– 3,000	73.1	11.3	15.6	41.7	6.2	52.2
3,000– 3,500	64.7	15.0	20.3	31.0	6.5	62.5
3,500– 4,000	59.6	16.4	24.0	24.7	7.4	67.9
4,000– 4,500	54.3	18.0	27.7	20.6	7.2	72.2
4,500– 5,000	52.4	18.8	28.8	17.3	7.2	75.5
5,000–10,000	40.4	19.8	39.8	9.3	6.1	84.6
10,000–15,000	30.9	18.5	50.6	4.9	4.6	90.5
15,000–20,000	26.8	19.3	53.9	3.7	4.2	92.1
20,000 & over	17.3	14.2	68.5	2.5	3.1	94.4

* Including imputed rent and profits.

† Excluding imputed income.

TABLE V

PER CENT OF CONSUMER UNITS RECEIVING INCOME ONLY FROM SPECIFIED SOURCES, UNITED STATES, 1936

Income class	Labor			Profits		Property		
	min.*	min.†	max.*	min.†	max.†	min.*	min.†	max.†
0- 500	28.5	61.7	65.7	3.7	3.7	0.2	15.8	15.8
500- 1,000	41.4	73.8	75.6	4.9	5.0	—	4.2	5.8
1,000- 1,500	34.9	66.1	69.1	4.5	5.2	—	1.7	3.4
1,500- 2,000	21.9	57.1	63.1	5.9	6.5	—	1.5	2.9
2,000- 2,500	2.9	42.0	45.8	6.7	7.1	—	2.0	3.0
2,500- 3,000	—	42.8	43.7	6.3	8.4	—	2.0	4.1
3,000- 3,500	—	29.9	35.7	5.3	8.5	—	3.0	4.4
3,500- 4,000	—	24.5	24.5	5.3	9.1	—	3.0	6.3
4,000- 4,500	—	17.1	21.9	5.7	9.4	—	4.1	6.9
4,500- 5,000	—	17.1	17.8	5.7	9.2	—	4.1	7.8
5,000-10,000	—	5.4	10.8	3.0	8.1	—	5.6	10.4
10,000-15,000	—	—	3.5	.7	5.5	1.6	8.5	16.0
15,000-20,000	—	—	3.1	—	4.2	1.5	8.4	15.4
20,000 & over	—	—	0.2	.9	2.3	4.6	11.5	16.7

* Allowing for imputed rent and profits. The minimum for profits, on this basis, is zero in all classes.

† Cash income only.

business cash profits and partnership profits, as shown by Table I, were proportionally most numerous in the classes above $3,500. But items of imputed profits (food, fuel and merchandise from stock) were commonest in the classes below $3,500. The ratio of all profits to total income, as reported on tax returns, varied slightly among the income classes, and was apparently highest between $3,000 and $15,000. According to the NRC, however, independent professional and business families (including farm families) were most frequent in the class below $500 and the classes over $4,000. Families were more likely than single individuals to receive profits, perhaps on account of the high proportion of married men among farmers. Single men also surpassed single women in the matter of profits, in every income class.

Only a small proportion of consumer units relied wholly on profits. That proportion was greatest in the classes between $1,500 and $4,000, and consisted mainly of families, although according to the Delaware returns a fairly high proportion of single men in the classes between $2,500 and $15,000 may have had no other kind of cash income.

PROPERTY INCOME

Property income is classified as dividends, interest, rent (including a small amount of royalties), and fiduciary income (i.e., income from estates and trusts). These will be taken up in that order, after which property owners of all sorts will be combined.

Dividends.—The total number of consumer units receiving dividends in 1936 was apparently about 4,731,000. Of course there were many other stockholders who received no dividends and in many families there was more than one stockholder. The total number of voters receiving dividends was over 8,200,000, allowing two to a family and assuming that each individual taxpayer is a voter. Dividend receivers were least numerous, proportionately, in the $500–$1,000 class, where only one person in nineteen received dividends; most numerous in the wealthiest class, where over 87 per cent received them. About half of the dividend receivers had total incomes from all sources amounting to less than $2,000, and they received about 10 per cent of the dividends. Of all kinds of income studied dividends were the least evenly distributed. Half of them were received by persons with incomes over $16,000. In the classes under $15,000 single women were much more likely to receive dividends than were single men, and families were less likely.

Interest.—A surprisingly large number of persons in the lowest income brackets, and of single women in all income brackets, received interest. Half of the interest receivers had total incomes under $1,615, and half of the interest income was received by consumer units with total incomes under $3,280. But if the survey made by the Consumer Purchases Study can be relied on there were apparently also a large number of families who had savings accounts, presumably paying interest, but who did not admit receiving any interest, either to the CPS investigators or on their tax returns. However, without counting these owners of non-interest paying savings bank accounts, or attempting to explain them, there were 4,434,000 families and 2,086,000 single individuals who received interest, or over 11,000,000 voters. Including the savings accounts there were 6,282,000 families, making the number of voters 14,650,000. The total amount received appears to have been $2,463,000,000.

Rents.—Rent is a form of income that is nearly as evenly distributed as wages and salaries, if the imputed rent received by home owners is included. There were in 1930, according to the Census, 14,002,000 home owners; in 1940 there were 15,196,000. Only 11,415,000 home-owning families can be located by the methods used in the present study, using the ratios of home ownership in each income class reported by the Consumer Purchases Survey. Half of them had incomes under $1,570. The tax reports enable us to locate 3,282,000 consumer units receiving cash rent, most of whom were probably also home owners, and all of whom have been treated as home owners in this analysis. But since some of them were not, the total number receiving rent, cash and imputed together, is underestimated. Cash rent was more evenly distributed than dividends, but not as evenly as interest. Half of the consumer units receiving it had incomes under $1,760, and half of the cash rent received was received by consumer units with incomes under $2,250.

Fiduciary (*Estates and Trusts*).—Fiduciary income was unimportant in amount and received by only a very small proportion of persons in each income class except the highest. In this study it has been necessary to lump fiduciary income together with some miscellaneous sources, such as annuities and other items partaking of the nature of interest but not reported as such on the tax returns.

All Property Income Combined.—When property incomes are lumped together the number in the 0–$500 class, and the amount, are surprisingly large. Checking the Delaware tax data against the CPS figures for families we find that the tax returns show a larger amount of interest and dividends than the CPS; in fact, they show a larger amount of either interest or dividends than the sum of the two as reported by the CPS. But there is reason to believe that the CPS proportions were below the truth.[2]

In the case of rent, other than imputed rent from owned homes, the CPS ratio of families receiving it was higher than the tax return ratio for classes between $500 and $1,500, although the ratio of rent income to total income was lower.

The CPS also reported an item of income other than earnings or rent or dividends or interest or imputed rent. This consisted mainly of annuities and pensions, which are of course property income, although of a non-transferable kind, since they are not paid for current services. Adding together all the property income reported by the CPS the total is a larger proportion of total income than that derived from the tax returns in each class under $2,000; and like the proportions derived from the tax returns they become larger as the income declines from $1,250 toward zero.

A similar increase is shown by the CPS study in the proportion of families receiving cash rent or interest and dividends or imputed rent, as the total income declines below $750. A little reflection shows why this must be so: a large proportion of the people in the lowest income class are there because they are incapable of earning anything—because of age or physical disability. Such small receipts as they may enjoy from property they have saved or inherited constitute a large proportion of their small incomes. In many cases they are able to spend more than their income by gradually drawing down their capital.

The CPS figures for imputed income from home ownership are included in property income in these tabulations. All of the tax return ratios used are based on total income, excluding capital gains, and where necessary further adjusted to allow for imputed income, which was taken from the NRC reports.

---

[2] See footnote 1.

THE PROLETARIAT AND THE PROPERTIED CLASS

The simple picture of a class society constructed by Karl Marx and accepted by unobserving persons in this country is not supported by the income tax statistics. We do not have a large propertyless proletariat, living in poverty and despair, ruled by a small class of capitalists, living in luxury and idleness. Every income class received over one-fifth of its income from sources other than wages; no income class received less than one-fifth of its income from the sale of its labor to employers; and no income class received less than 37 per cent of its income from labor of some sort—either working for others or actively conducting a business enterprise. The number of idle rich is small. Of the 231,000 consumer units with incomes over $10,000 not over 51,000 lived entirely on income from property and at least 16,000 or 17,000 of these were single women (including widows).

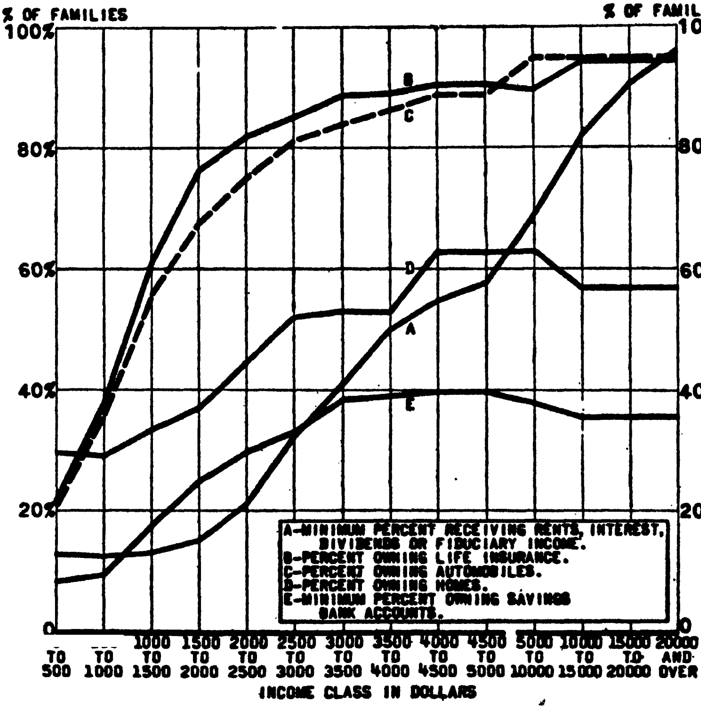
Of the 5,915,000 consumer units (if that is the right number) with incomes under $500 on the other hand, 935,000 were entirely dependent on property for their income, except if they were able to raise part of their own food and fuel; and of these some 370,000 were single women. One-half of the persons who received any income from property had total incomes under $1,110. The median income of all consumer units receiving wages or salaries was $1,220, while the median income from all sources was $1,214. Another important fact, not shown by the statistics but known to all of us, is that most of the male capitalists had earned their capital by their own efforts, and many of the propertyless laborers will earn enough capital, before they become too old to work, to maintain them in their old age. As a corollary of this we find a large proportion of persons in the lowest 10 per cent of the population—the "submerged tenth"—who live entirely or mainly on the fruits of their savings. These, and the single women of all income classes, constituted almost the whole of our non-working "leisure class" of capitalists.

The corporations of the country were not owned by sixty families or 200 large stockholders. It required at least 126,000 of the wealthiest consumer units to muster one-half of the shares. Half of the stockholding units had incomes under $2,030. The unincorporated businesses of the country were even more widely owned; half of the 9,800,000 consumer units receiving business or partnership profits (omitting imputed profits) had incomes under $1,295.

The homes of the country were owned by about 15,000,000 consumer units, not counting the number indirectly owned through corporations and associations. There were at least 12,700,000 consumer units who owned bonds, mortgages or bank accounts, and half of them had incomes under $1,780.

Some other important forms of capital were not referred to in the tax returns. Some information is available from the Consumer Purchases

PROPORTION OF FAMILIES OWNING PROPERTY
BY INCOME CLASSES
(United States-1936)



Study concerning the ownership of automobiles and insurance policies, and is shown in the attached chart. Clothing and furniture are of course very widely owned but may be regarded as consumer goods rather than capital. But the ownership of a most important type of capital—the tools and machines, shops, stores, inventories and farm equipment necessary to earn business or partnership profits—can only be surmised from the distribution of profits.

On the whole the best measure of the distribution of ownership of all forms of productive capital together appears to be the combined income from property and profits. According to this measure one-half of all the productive capital of the country was owned by consumer units with incomes under $3,065, and half of the owners had incomes under $1,575.

We must conclude that the vast bulk of our people are both capitalists and laborers during the prime of their lives; they begin their careers without much capital except their bodies and their brains, and they wind up with very little earning power in them, but with enough capital to live on. This fact largely explains the political stability we have had in the past and is the best hope for political stability in the future. Certain other factors in our political stability cannot be measured by the available material, but should be covered by future investigations. They include the age and race factors—to what extent the propertyless persons are young, or Negro, or recently arrived foreigners—and the fluidity factor, by which is meant the shift of individuals from income class to income class and the length of life of large fortunes. My own belief is that, if those factors were studied it would be found that in the matters of equality of opportunity and wide distribution of wealth, the United States compares favorably with any nation now existing, or any important civilized nation of the past.

NOTE: Several tables showing in greater detail the facts mentioned in this article had to be omitted for lack of space. A limited number of copies will be sent by the author to those requesting them.

COMPARISON OF DEMOGRAPHIC RATES ASSUMED BY THE NATIONAL RESOURCES COMMITTEE WITH ACTUAL EXPERIENCE

BY ROBERT J. MYERS

IN ITS POPULATION projections, the National Resources Committee assumed certain death rates and birth rates for future periods[1] up to 1980. The assumed future rates had as a base point the actual rates for the period 1930–34. Now that five years have elapsed and actual census material is available, it is worth while to make a comparison of the actual rates with those estimated. It is hardly necessary to point out that forecasts of this type are especially dubious for the immediate future because of the effects of the war.

In order to facilitate computations, the type of death rate used by the NRC differed somewhat from the usual age-specific rates of life tables. For a given quinquennial age group the probability of dying within five years was set down for various future quinquennial periods. In order to project births, there was developed a special death rate to give the probability of uniformly-spaced births in a five-year period dying before the end of that time. Three sets of rates, low, medium, and high, were developed for native whites, foreign-born whites, and Negroes with, of course, separate rates for males and females.

Perhaps the best actual mortality data for current purposes is the set of 1940 population life tables[2] prepared by the Statistical Bureau of the Metropolitan Life Insurance Company,[3] from which death rates of the same form as used by the NRC may be developed. In order to make a consistent comparison, it is necessary to make certain adjustments in the NRC figures. First, NRC figures applicable to 1940 may be obtained by averaging the rates for any particular age group for the two quinquennial periods, 1935–39 and 1940–44. Second, a combination of the NRC mortality rates for native whites and foreign-born whites must be made since the Metropolitan life tables which are being used for comparison are only on the basis of total whites. This was accomplished by weighting native white and foreign-born white death rates by the respective proportions of these groups relative to the total white population as indicated by the 1940 Census.[4]

---

[1] The actual figures appear in Tables 1B and 1C of *Population Statistics*, Vol. 1, National Data, National Resources Committee, October 1937. Details of the methodology are contained in P. K. Whelpton, "An Empirical Method of Calculating Future Population," this JOURNAL, September 1936, pp. 457–473.

[2] It will be noted that the actual mortality experience is that for only a single year, 1940. However, the mortality of that year differed only slightly from that of surrounding years; the crude death rate for 1940 was 10.7 per thousand as contrasted with an average of 10.6 for the three years 1939–41.

[3] Presented in abridged form in *Statistical Bulletin*, Metropolitan Life Insurance Company, December 1941, p. 8. The complete table was made available through the courtesy of Dr. A. J. Lotka.

In Table I there are set down separately by sex and race the actual 1940 death rates on a comparable basis with the NRC figures. Table IIa and IIb compare these rates with those of the NRC modified as described previously to apply to 1940. At the younger ages the actual rates for both white males and white females are appreciably lower than even the "low" estimate; thus, for white males the differential is 20 per cent at the youngest childhood ages and prevails to some extent to age 35, whereas for white females the differential is even larger and persists until the older ages. For white males the actual rates beyond age 35 are about 5 per cent in excess of the "low" estimates, but for white females the corresponding estimates are about 5 per cent too high for ages 35

TABLE I

ACTUAL 1940 FIVE-YEAR TIME INTERVAL DEATH RATES
FOR QUINQUENNIAL AGE GROUPS*

Age group	White persons		Colored persons	
	Males	Females	Males	Females
Birth to 0–4†	48.5	38.1	75.9	59.6
0– 4	10.1	8.3	16.4	13.8
5– 9	5.1	3.7	7.0	5.9
10–14	6.7	4.7	13.1	13.9
15–19	9.8	6.9	25.0	25.0
20–24	12.0	8.9	34.9	31.1
25–29	13.8	10.8	41.8	36.5
30–34	17.5	13.7	49.9	43.8
35–39	24.6	18.1	63.6	57.7
40–44	36.7	25.5	85.3	76.4
45–49	55.4	36.9	117.9	103.9
50–54	82.5	54.1	150.7	136.7
55–59	118.7	80.7	171.2	155.6
60–64	169.5	124.9	209.9	183.0
65–69	243.1	193.8	265.6	230.6
70–74	345.1	289.5	326.5	288.4
75–79	476.5	426.0	412.6	349.0
80–84	623.9	600.0	546.2	420.4
85–89	761.5	774.9	699.7	555.7
90–94	864.9	882.6	832.1	732.6

* Derived from U. S. 1940 Life Tables prepared by the Statistical Bureau of the Metropolitan Life Insurance Company.

† This represents the number of deaths within a five-year time interval among 1,000 infants born uniformly throughout that period.

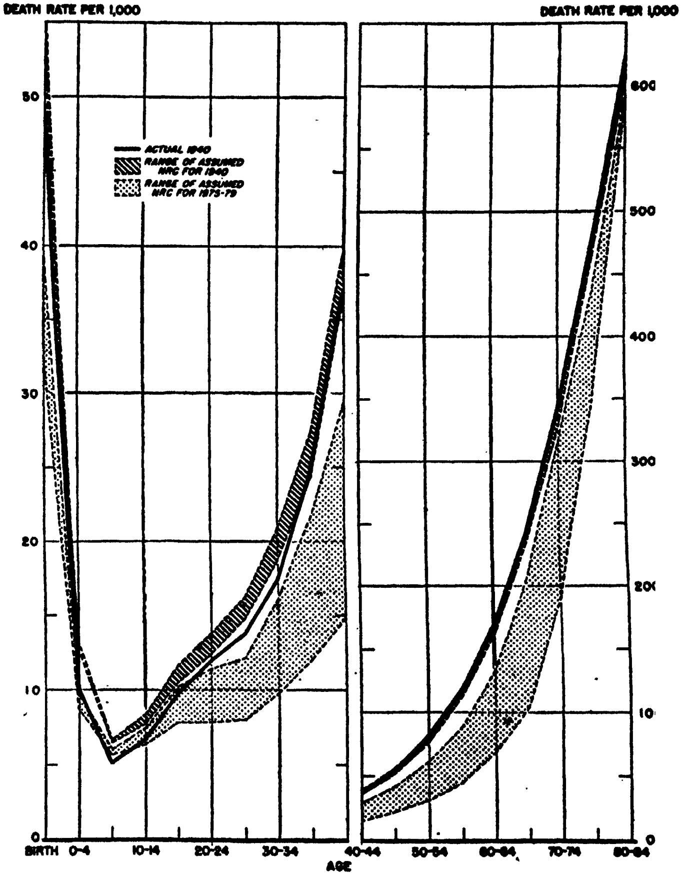
---

[4] Based on the release dated April 10, 1942, *Age Composition of the Population*, Series P-10, No. 6, Bureau of the Census.

through 55. Up to age 50 the "high" estimates exceed the actual rates for white males, but beyond there is a slight differential in the opposite direction. The "high" rates for white females are above the actual death rates at all except the extreme ages.

CHART I

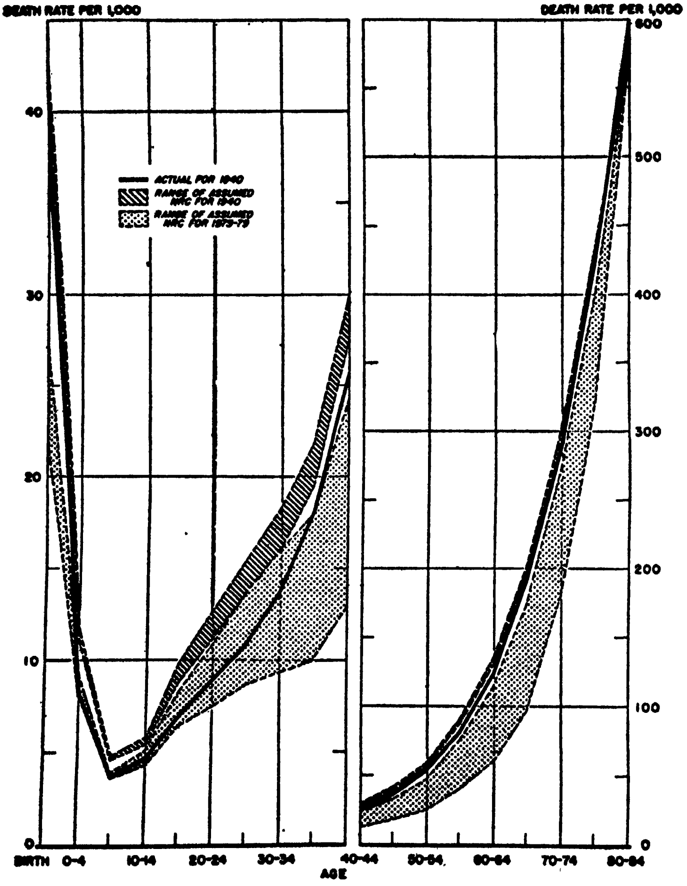
COMPARISON OF ACTUAL 1940 FIVE-YEAR DEATH RATES AND THOSE OF NRC—WHITE MALES



Charts I and II dealing with white males and white females respectively make a graphical comparison of the actual 1940 death rates with those indicated for 1940 by the NRC and also with those for the "ultimate" condition in 1975–79. The difference between the actual rates and the 1940 estimates at the younger ages is quite striking, but beyond age 35 for males and 55 for females there is little difference. At the childhood ages for both males and females and at the younger adult ages for females the actual 1940 rates fall well within and, in some instances, below the ultimate range which, it might be pointed out was extremely narrow at the younger ages, although wisely wide with advancing age.

CHART II

COMPARISON OF ACTUAL 1940 FIVE-YEAR DEATH RATES AND THOSE OF NRC—WHITE FEMALES



Actual 1940 colored death rates have fallen fairly closely within the range developed by the low and high estimate as may be seen from Table IIb. The greater amount of fluctuations present as contrasted with the fairly smooth trend for white persons in Table IIa arises from the well-known difficulties in obtaining accurate reporting for this population group.

TABLE II

ACTUAL 1940 FIVE-YEAR TIME INTERVAL DEATH RATES AS PERCENTAGE OF ASSUMED NATIONAL RESOURCES COMMITTEE RATES

Age group	a. White persons*				b. Colored persons†			
	Males		Females		Males		Females	
	as per cent of				as per cent of			
	"Low"	"High"	"Low"	"High"	"Low"	"High"	"Low"	"High"
Birth to 0–4‡	94	89	95	90	111	102	107	99
0– 4	80	77	75	72	96	88	89	80
5– 9	80	77	80	77	90	80	95	84
10–14	87	81	87	81	109	91	122	98
15–19	94	84	78	70	107	86	102	81
20–24	97	87	79	70	108	89	108	87
25–29	93	85	79	70	106	91	104	87
30–34	94	84	85	75	98	84	104	87
35–39	101	90	93	83	104	90	108	93
40–44	101	93	93	85	103	93	105	94
45–49	107	99	95	89	114	105	111	102
50–54	109	103	96	91	117	110	105	98
55–59	105	102	92	89	105	101	94	90
60–64	105	101	96	92	103	100	91	87
65–69	104	101	99	96	102	99	95	92
70–74	105	101	99	95	103	99	102	98
75–79	103	101	100	98	94	92	95	94
80–84	103	102	102	102	95	95	97	97
85–89	104	104	107	107	97	97	106	106
90–94	104	104	109	109	104	104	109	109

* Averages of 1935–39 and 1940–44 rates for native and foreign-born whites weighted by respective proportions in 1940 census for given age-sex group (for quinquennial age groups beyond age 75, where census data are not subdivided by age groups, proportion for those 75 and over is used).

† Averages of 1935–39 and 1940–44 rates for Negroes.

‡ See footnote †, Table I.

One other comparison that may be made is to determine the calendar year in which the assumed "medium" death rate is the same as the actual 1940 one (Table III). This will indicate the length of time that has been "gained" in regard to "expected" mortality improvement. For white males up to age 40 there has been a gain of about 5–10 years with the exception of ages 5–15 where we are now already lower than the ultimate figure. For white females there is even more of a gain with such periods being 20–40 years up to age 30 but less thereafter up to age 60. For colored persons there are relatively few gains, except that at the older ages actual 1940 death rates are in many instances lower than the estimated ones; how much validity may be attached thereto is open to question.

TABLE III

CALENDAR YEAR IN WHICH ASSUMED NATIONAL RESOURCES COMMITTEE "MEDIUM" DEATH RATE EQUALS ACTUAL 1940 RATE*

Age group	White persons		Colored persons	
	Male	Female	Male	Female
Birth to 0–4†	1944	1942	‡	‡
0– 4	1951	Ultimate§	1942	1947
5– 9	Ultimate§	1972	1970	1943
10–14	Ultimate§	1962	‡	‡
15–19	1947	Ultimate§	‡	1942
20–24	1944	1962	‡	‡
25–29	1944	1967	‡	‡
30–34	1944	1955	1942	1941
35–39	1941	1947	1941	‡
40–44	‡	1947	1941	‡
45–49	‡	1945	‡	‡
50–54	‡	1941	‡	‡
55–59	‡	1941	‡	¶
60–64	‡	‡	‡¶	¶

* Rates for quinquennial calendar periods assumed to apply to middle year, with linear interpolation when actual rate falls between assumed ones. Native white rates used for white persons and Negro rates for colored persons.

† See footnote †, Table I.

‡ Actual 1940 rate higher than assumed 1940 "medium" rate derived as described in footnote * of Table IIa and IIb.

§ Actual 1940 rate lower than assumed ultimate (1975–79) "medium" rate.

¶ At most of the ages beyond 55 actual colored death rates are lower than the NRC "medium" estimate for 1940 (see Table II).

The derivation of the actual five-year time interval birth rates was, in general, performed exactly according to the definition which relates births in the given quinquennium (1935–39) to women surviving at the end of the period according to what the age of the mother would have been at that time (regardless of whether or not she actually was alive then). In order to compute such rates, there was first obtained the distribution by single years of age of mother and by race (white and colored) for reported births in each of the years 1935–39.[5]

The births for each of the five years were then collected on a "step" basis so that there was obtained the number of births in the quinquennium to women at each given single age in 1939. For example, the figure for women age 20 in 1939 included births to women age 20 in 1939, age 19 in 1938, etc. Since the distribution of the total female population is known only as of the census data, April 1940, and since the births were related in effect to ages of women as of the middle of 1939, a further adjustment was necessary. Linear interpolation between the births at single years of ages was performed to make this ¾ year adjustment. In other words, it was assumed that ¾ of the births to women aged x in 1939 applied to women aged $x+1$ as of the census data in 1940, while the remaining ¼ applied to women aged x then. Following this subdivision, quinquennial totals were obtained and divided by the 1940 population figures to produce the desired birth rates on a basis unadjusted for under-reporting of births. Such adjustment was then made on the basis used by Whelpton;[1] namely, an under-reporting factor of 96 per cent for white births and 86 per cent for colored births. If reporting accuracy has improved, the finally resulting "actual" birth rates are overstated relative to the NRC assumed rates, but as yet there seems to be no specific evidence on this point.

---

[5] From various issues of *Vital Statistics of the United States*, Bureau of the Census. A slight approximation had to be introduced since single age data are not available for Massachusetts and New Hampshire, but rather only quinquennial values; it was assumed that the distribution within each quinquennial age group was the same for these two states as for the remainder of the country, a procedure which can produce only a negligible error.

Table IV shows the actual 1935–39 rates for each quinquennial age group from 10–14 to 55–59, while Chart III deals with the range 15–19 to 45–49 since the NRC did not present figures for the youngest quinquennial age group nor for the two oldest ones, probably because of their small magnitudes. In addition, these actual rates are shown as percentages of the low and high NRC assumptions, with their rates for native and foreign-born whites being blended as was done for death rates and likewise the Negro rates being assumed to apply to colored persons. The total of the actual rates for all ages is indicative of the gross reproduction rate, although not an accurate expression thereof. For white females the actual rates tend to be higher than the low assumptions except for the youngest and oldest age groups but, with only one exception are lower than the high assumptions. Thus at age 25-29 the actual rate is almost 20 per cent in excess of the low assumption and slightly greater than the high assumption. In the aggregate, the actual birth rates are about 10 per cent greater than the low assumptions and slightly lower than the high assumptions. It will be noted that in Chart III there is a considerable margin between the actual 1935–39 rates and the low assumption for 1975–79 (the high assumption for that period being the same as the high assumption for 1935–39).

The actual birth rates for colored females, when contrasted with the NRC assumptions in Table IV, do not show the same smooth relative trends as were present for white females. As with mortality rates, the greater fluctuation undoubtedly arises from inaccuracies and incompleteness of reporting. In the aggregate, the same trends are shown as for white females; namely, an excess of about 10 per cent over the low assumptions and a slight deficiency as compared with the high assumptions.

In summary, it may be said that our actual mortality rates have improved appreciably more than the National Resources Committee estimated for the younger ages, whereas at the older ages there appears to be some indication of less improvement, especially for white males. In any future projections it would seem advisable to introduce an even more drastic reduction in mortality rates at ages up to 35 (disregarding the immediate and near-future effects of the war, particularly upon males).In this connection it may be noted that the 1941 Hypothetical Life Table,[6] which indicates the plausible long-range future level of mortality and was developed from a corresponding table prepared in 1933, made changes in mortality rates only at ages under 50.

TABLE IV

ACTUAL 1935–39 FIVE-YEAR TIME INTERVAL BIRTH RATES FOR QUINQUENNIAL AGE GROUPS* COMPARED WITH ASSUMED NATIONAL RESOURCES COMMITTEE RATES†

Age group	White females			Colored females		
	Actual rate	As per cent of "Low"	As per cent of "High"	Actual rate	As per cent of "Low"	As per cent of "High"
10–14	.1	‡	‡	1.6	‡	‡
15–19	55.1	91	85	201.3	129	115
20–24	430.7	106	99	716.3	101	91
25–29	616.7	117	103	675.7	121	106
30–34	496.8	110	97	492.4	118	105
35–39	312.6	104	90	304.8	87	75
40–44	163.9	100	87	209.3	118	101
45–49	42.6	97	84	55.5	97	82
50–54	3.1	‡	‡	8.3	‡	‡
55–59	.1	‡	‡	.6	‡	‡
Total	2121.7	109	97	2665.8	110	97

* Derived from recorded birth statistics for 1935–39 and the 1940 census as described in text.

† Native and foreign-born white rates weighted by respective proportions in 1940 census to give total white rates. Negro rates assumed to apply to total colored persons.

‡ National Resources Committee did not present figures for this age group.

Our actual birth rates have been slightly higher than the medium assumptions of the National Resources Committee but below their high assumptions. As for the main child-producing period, 20–34, the actual birth rates have been about as large as the high assumptions, but at the older ages this tendency was not maintained. A more detailed analysis by child-order would show that this trend arises from the decline in the birth rate for third and higher children.

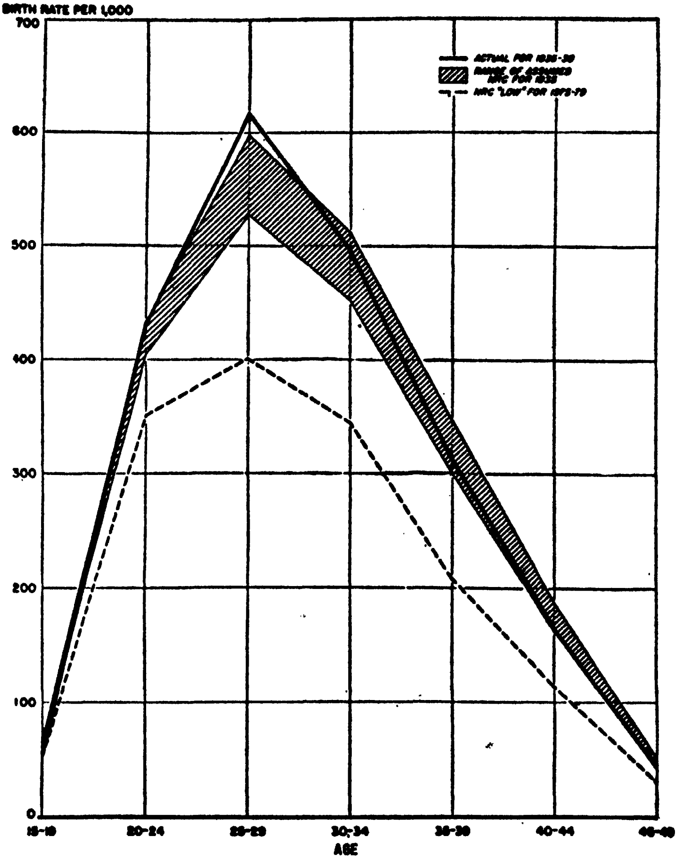
The drastic increase in the birth rate since 1939 would, of course, raise the actual rates considerably above the National Resources Committee assumptions which were based on the birth rates in the trough of the depression, 1930–34. Most authorities are prone to believe that

---

[6] *Statistical Bulletin*, Metropolitan Life Insurance Company, November 1941, p. 3.

CHART III

COMPARISON OF ACTUAL 1935–39 FIVE-YEAR BIRTH RATES AND THOSE OF NRC—WHITE FEMALES



the present increase is purely a temporary phenomenon and that after the war age-specific birth rates will continue to decline as is provided in the NRC medium and high assumptions. It is not at all impossible that the present fertility state of the nation is merely a forerunner of what might be expected after the war when economic reconstruction along the lines of the Atlantic Charter will produce social conditions favorable to entirely higher fertility trends.

CHOICE OF THE DEPENDENT VARIABLE IN REGRESSION ANALYSIS

By Frederick V. Waugh
Food Distribution Administration

In the simplest problem of two variables it is always possible to calculate two linear regressions: one with X_1 as the dependent variable, and one with X_2 as the dependent variable. Unless there is perfect correlation between the observed values of X_1 and X_2 these two regression equations will be different. If we plot both on a diagram they will cross each other. How do we know, in any particular case, which regression equation to use?

In a multiple regression problem with n variables we can always calculate n different regression equations treating each variable, in turn, as dependent and the other $(n-1)$ variables as independent. How do we know, in multiple regression, which of these n equations to use?

There appears to be a good deal of confusion about this rather simple matter. A common rule of thumb procedure—particularly among economists—is to select as the dependent variable the one which is taken to represent the "result" or "effect," while the independent variable, or variables, are taken to represent "causes." Those who follow this procedure often argue that only one of the n regressions has any real meaning or any use.

This procedure finds support in some of the best and most recent textbooks. To illustrate this, we shall make several references to the recently revised edition of Dr. Ezekiel's excellent book.[1] Similar references might be found in other textbooks.[2]

Dr. Ezekiel[3] writes,

> It is always necessary to think over the problem carefully, and decide which variable we are going to regard as the independent or causal variable and which one as the dependent, or resultant. Thus, if we were relating variations in tobacco yields to applications of fertilizer, obviously the differences in fertilizer would be the cause and differences in yield the result, so we would sort our records according to the differences in fertilizer.

The last clause might just as well have been written, "so we must use yields as the dependent variable." Dr. Ezekiel states here that we should *always* choose the dependent variable from considerations of cause and effect, but in the following paragraph he indicates that this procedure is not always possible. Thus, he writes:[4]

---

[1] Mordecai Ezekiel, *Methods of Correlation Analysis*, 2nd edition, John Wiley & Sons, New York, 1941.

[2] For example, Albert E. Waugh, *Elements of Statistical Method*, McGraw-Hill Book Company, New York, 1938. See headings, "Nature of Multiple Relationships" and "Dependent and Independent Variables," on pps. 318–319.

[3] Ezekiel, *op. cit.*, pp. 50–51.

> It is frequently impossible to say which variable is the cause and which is the effect. All that can be definitely established is that the two vary together. Yet one may wish to regard one variable as the one whose values are given or known. It is then called the independent variable. . . . The second variable will then be regarded as the one whose values are to be related to, or estimated from, the values of the known variable. It is then called the dependent variable since it is treated as depending upon the given values of the independent variable.

These two quotations appear to summarize Dr. Ezekiel's views on this problem. He seems to advise us to follow the cause and effect principle whenever we know which is which. When we don't know which is cause and which is effect, or when both variables are results of a common cause, we are advised to use another criterion in choosing the dependent variable. Not only does Dr. Ezekiel advise us to follow this procedure, he follows it consistently himself in all the problems presented in his book. Whenever one variable is clearly the result, it is chosen as the dependent, and in such cases Dr. Ezekiel seems to feel that only this one regression equation has any meaning or any use. It is largely on these grounds that Dr. Ezekiel recommends in his technical appendix[5] a computational method which gives only one of the n multiple regression equations, when it would be possible with very little additional work to get them all.[6]

What about this procedure? Does it give us the regression equation we want in each particular case?

Suppose we take the example given by Dr. Ezekiel: tobacco yields and fertilizer applications. Let us imagine that we have surveyed 100 farms in North Carolina, and have found the average yield of tobacco and the rate of fertilizer application on each farm. We want a line or equation which will express the relation between the two variables: tobacco yield and fertilizer application. There are two such lines (neglecting the possibility of choosing between different types of curves)—one with yield dependent and one with fertilizer dependent. Dr. Ezekiel says that variations in fertilizer are obviously to be considered as the cause of the variations in yield, and, therefore, that yield must be the dependent variable. Granting that the direction of the cause and effect relation is clear in this case, the writer fails to see any reason for con-

[4] *Op. cit.*, p. 51.

[5] *Op. cit.*, Appendix I, pp. 455–485.

[6] Frederick V. Waugh, "A Simplified Method of Determining Multiple Regression Analysis," this JOURNAL, December 1935, and "The Analysis of Regression in Subsets of Variables," this JOURNAL, December 1936.

cluding that yield must be the dependent variable. We might use this set of data to answer either of the following questions:

a. What were the average yields of tobacco associated with various amounts of fertilizer? The answer to this question is provided by the regression with yield as the dependent variable.

b. What were the average amounts of fertilizer associated with various yields of tobacco? The answer to this question is provided by the regression with fertilizer application as the dependent variable.

We shall not prove these statements here. The proofs can be found in practically any textbook on correlation and regression.[7] The point is that *both* regressions have definite meanings and definite uses, and our choice does not depend at all upon which is cause and which is effect, but simply upon the question we want answered. We may very well want both regressions, because we may very well want to answer both questions.

The same point may be illustrated with any problem which may come to the reader's mind. However clear the direction of the cause and effect relation may be there are always two essentially different questions which may be asked in any problem of two variables. If we want the average, or expected, value of X_1 associated with any particular value of X_2, the answer is given by the regression with X_1 dependent—regardless of the nature of cause and effect. If we want the average, or expected, value of X_2 associated with any particular value of X_1, the answer is given by the regression with X_2 dependent—regardless of the nature of cause and effect.

If we want to estimate the average, or expected, height of fathers corresponding to a given height of sons, we should take height of fathers as the dependent, regardless of the fact that this must be the causal factor. If we want the average amount of potato production associated with parity prices of potatoes, we should take potato production as the dependent, even if we are sure that variations in production *cause* the variations in price.

In a big multiple regression problem we often think in terms of one particular regression. Thus, in a farm management survey we may be concerned with a large number of (causal) factors which affect farm income (the resultant). In such a problem our thinking tends usually to run in the direction of investigating what changes in average, or expected, income would tend to be associated with particular adjustments in the several causal variables. We tend to frame our questions in such terms as, "How much increase in income could be expected on this farm if we added 50 chickens?" or "If these farmers fed each of their cows another pound of feed each day, how much would their average income be increased?" Both of these questions are answered best by the regression with income as the dependent variable.

---

[7] For example, in Ezekiel, Technical Note 2, on page 489. Note that Y is selected as the dependent variable without any consideration of whether it represents cause or effect. In any case the regression minimizes the sum of squares of the Y residuals. And in any case a similar mathematical development would show that by choosing X as the dependent variable we minimize the sum of squares of the X residuals.

Does this mean that the other $(n-1)$ regressions have no value? Not at all. Let us think of our problem in different terms. Let us suppose that some of the farmers in our survey must increase their incomes by $200 a year to get a minimum standard of living. They hope to do this by raising more chickens, but without making any other adjustments. We want to know then the average number of chickens associated with the higher income and with the same quantities of the other factors. The answer to this is given by the regression with the number of chickens as the dependent, and with income and all the other variables as independents.

Further examples should be unnecessary. The reader can easily apply similar reasoning to *any problem* of related variables, and will find that in each case there will be as many useful regression equations as variables, and that each of these equations provides the answer to a reasonable question.

Now, to summarize, and to add a few general comments.

1. In any problem of relationships between n variables, each of the n regression equations has a definite meaning, and each can be used to answer a specific type of question.

2. Even when our problem is such that one variable is clearly the "result," and the other variables are clearly "causes," we must choose one of the "causal" variables as the dependent if we wish to estimate the average, or expected, value of that variable corresponding with any set of values for the others.

3. We may very often wish to estimate the average, or expected, value of one of the "causal" variables needed to produce a stated "result." In any such case the "causal" variable becomes the dependent. The "resultant" variable, and any other "causal" variables are treated as independents.

4. There is no sound reason for the common rule of thumb practice of always designating current price as the dependent when relating it to such variables as current supplies and current incomes of consumers; nor for always designating consumption as the dependent when it is related to current retail prices; nor for always designating acreage of a crop as the dependent variable when it is related to prices of that crop; nor for always designating the yield of a crop as the dependent when re-

lating it to fertilizer applications and other variables representing cultural practices. This rule of thumb practice is based on the mistaken idea that the only useful regression equation is the one with the "resultant" variable as the dependent.

5. Nothing in this paper implies, in any way or in any degree, that we can neglect careful logical analysis in the statistical measurement of relationships. The writer subscribes completely to Ezekiel's chapter[8] emphasizing the need for such logical analysis. The point is only that considerations of cause and effect do not indicate which variable should be dependent.

6. We are concerned in this paper only with the classical regression equations. These are to be used whenever our purpose is to estimate values of one of the observed variables. If, however, we assume perfect correlation between the true, underlying values of our variables, we are assuming that departures are due to errors of observation, or to what Frisch[9] calls "disturbances." If we can estimate the relative magnitudes of these errors or disturbances in each of the variables it is possible to calculate one line in the case of two variables, or one n-dimensional plane in the case of n variables expressing the relationship between the true, or basic values. In this case there is only one line or plane, and no one variable is taken as dependent. The so-called "orthogonal regression" assumes equal errors, or disturbances, in each variable. No such line or plane gives any measure of correlation, since correlation is assumed to be perfect among the true, underlying values. Neither does such a line or plane provide an estimate of the average, or expected, values of any of the observed variables.

---

[8] *Op. cit.*, Chapter 24, pp. 442–453.

[9] Ragnar Frisch, "Statistical Confluence Analysis by Means of Complete Regression Systems," Universitets Økonomiske Institutt, Oslo, Norway, 1934, especially Part II, pp. 47–83.

COMMENTS

BY MORDECAI EZEKIEL
United States Department of Agriculture

Dr. Waugh agrees that where we wish to estimate the most probable value of one variable from a given value of another, we take the latter as the independent variable, and the former as the dependent. He then goes on to say, however, that "both regressions have definite meanings and definite uses, and our choice does not depend at all upon which is cause and which is effect, but simply upon the question which we want answered." My own belief is that in cases where we have definite nonstatistical knowledge of which variable is cause and which is the effect, the reverse regression, which measures the average or most probable value of the "cause" associated with given values of the "result," is generally of little significance and is likely to be used for illogical or incorrect conclusions. Where, by the use of multiple correlation, we can eliminate other associated causes and come still closer to a measurement of the influence of each separate cause upon the result, the use of the reverse partial regression is, I believe, likely to be even more misleading.

The misleading character of the reverse regression is shown by Dr. Waugh's own illustration of farmers who wish to increase their incomes by $200 a year by raising more chickens, while holding the rest of their farming unchanged. Dr. Waugh would apparently determine this by finding the net regression of chicken numbers on farm income, while holding constant the other independent variables. Yet, for each individual farmer involved, what he needs to know is not "What numbers of chickens, on the average, are associated with given incomes?" but rather "How many chickens need I add to my farm to increase my income $200?" What the farmer can do is to *change the number of chickens;* that is the *cause* which he will employ to change his income. The most accurate way to find how much *each chicken* adds to income is to employ the net regression of income upon number of chickens. If this comes out 50 cents, that will mean the 400 chickens will be the number probably required. If the question is restated as how many chickens for a given sum instead of how much money per chicken, that does not change the fact that the *chickens* are what the farmer can actually change.

This differs from Dr. Waugh's point of view. He states "we may wish to estimate the average, or expected, value of that variable (chickens) corresponding with any set of values for the others (income)." While "we" may at times wish to know how many hens the $500 farmer had, and how many the $700 farmer had, the farmer has only an academic interest in those facts. What he wants to know is, how much effect will a *change* in hens have upon his income, for the *hens* are the only thing he can vary directly. The income which will follow as a consequence is still unknown when the hen numbers are decided upon.

I do agree that if we were to take new sample farms from the same universe, and to calculate for those new samples the average number of hens on farms having given incomes, the best advance estimate we could make of that relation would be based on the regression of hens on income in our original sample. But I question whether that information is of any value to the farmer who has to decide what he will do with his poultry flock—or for that matter, is often of very much value to anyone else.

There has been much fallacious reasoning based upon the reverse regression approach. One common illustration is the type of farm management report which shows the average number of acres of each crop and head of each class of livestock on farms *reporting various levels of income.* It has been the evident conclusion that because the men who had high incomes had more dairy cows, it would therefore pay the men who had lower incomes to increase the number of their cows. Yet actually a classification of the same data by dairy cow numbers might show only insignificant changes in average incomes with differences in numbers of cows. Similarly, in the poultry case, because of some other variable not explicitly measured—such as the ability of the more intelligent farmers to handle all their livestock better—the data might show relatively small increases in chickens associated with very large differences in income, perhaps one more hen for each ten dollars more income. Dr. Waugh would then conclude that a small change in chicken numbers was associated with a corresponding large increase in income. This information would tend to mislead the individual farmers, who after all were in position to change their chicken numbers but not to change their own inherent abilities. In fact, the lower the correlation between income and hens, and the smaller the differences in hens associated with given differences in incomes, the *fewer* the number of hens which this method would indicate as needed to increase farm income by $200!

In view of these considerations, I still believe that the procedures recommended in my book are the most satisfactory for the practical research worker who expects to use his research as a basis upon which to advise on the probable consequences of actions in the real world, for in that world, people need to know what effects (results) are most likely to follow from given things they may select to do (causes).

NOTE ON A COMMON STATISTICAL INEQUALITY

By Irving H. Siegel
U. S. Bureau of Labor Statistics

THIS PAPER is devoted to the discussion of a family of inequalities frequently encountered in statistical work. The prototype relation is

$$\sum st \sum uv \gtreqless \sum sv \sum ut \tag{1}$$

or

$$\Delta = \begin{vmatrix} \sum st & \sum sv \\ \sum ut & \sum uv \end{vmatrix} \gtreqless 0, \tag{1a}$$

where all of the sums include the same number of items. In the terminology of Hardy-Littlewood-Pólya, the inequality is "homogeneous in $\sum$" as well as in the variables, and the two sides of (1) are not necessarily "comparable." That is, the sense of the inequality is not affected by the substitution of means for the corresponding sums, but it need not be fixed for all non-negative values of the variables.[1]

In the following, we shall first present some useful identities related to (1a) and then analyze some cases of the inequality which are of interest to the statistician. The validity of the identities is independent of the sign of the variables. In some of the inequalities mentioned later, however, one or more of the variables will be restricted to non-negative values. In any case, only real discrete variables are considered; and all of the sums are finite. Analogous inequalities generally exist for integrals and infinite sums.

The factorability of the matrix of Δ into two rectangular ones is a very useful property. If each of the sums contains n elements, we have

$$\left\| \begin{matrix} \sum st & \sum sv \\ \sum ut & \sum uv \end{matrix} \right\| = \left\| \begin{matrix} s_1 \cdots s_n \\ u_1 \cdots u_n \end{matrix} \right\| \cdot \left\| \begin{matrix} t_1 \cdots t_n \\ v_1 \cdots v_n \end{matrix} \right\| \tag{2}$$

For our purposes, the rectangular matrices may be considered permutable: the multiplication is performed row by row, and the value of Δ is not altered by the interchange of diagonal terms. The value of Δ also remains unaltered if the order of the rows is reversed in both rectangular matrices, but its sign is changed if the order is reversed in only one. The matrix of Δ may also be decomposed into products of other rectangular matrices; for the sums in Δ may be expressed in terms of new variables (derived from the original ones) without being altered in value or general form.

---

[1] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities* (1934), pp. 1–6.

By virtue of (2) or any of its variants, Δ may be transformed into a double sum containing the ${}_nC_2$ products of second-order determinants obtainable from the rectangular matrices. For example, it follows from (2) that

$$\Delta = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \begin{matrix} s_i & s_j \\ u_i & u_j \end{matrix} \quad \begin{matrix} t_i & t_j \\ v_i & v_j \end{matrix} \tag{3}$$

In many special cases of importance, each determinant product is positive (or negative) and Δ is necessarily greater (or less) than zero.

The sum (3) may be regarded as the Cartesian form of a well known vector identity which is often associated with the name of Lagrange. If the s_i, t_i, u_i, and $v_i (i=1, \cdots, n)$ are the n rectangular coordinates of the vectors $\boldsymbol{s}$, $\boldsymbol{t}$, $\boldsymbol{u}$, and $\boldsymbol{v}$, respectively, then

$$\Delta = \begin{matrix} s\cdot t & s\cdot v \\ u\cdot t & u\cdot v \end{matrix} = (s \times u)\cdot(t \times v), \tag{4}$$

where a dot signifies a scalar product and a cross signifies a vector product.[2] As will be illustrated later, it is sometimes enlightening to translate (4) into its trigonometric equivalent.

There is often some advantage to be gained by the introduction of a weighted correlation coefficient which always has either the same sign as Δ or exactly the opposite.[3] In either case, the correlation coefficient and Δ vanish identically. For example, we may write

$$\begin{aligned} \Delta &= \sum st\left(\sum uv - \frac{\sum ut \sum sv}{\sum st}\right) \\ &= \sum st\left[\sum st\left(u/s - \frac{\sum stu/s}{\sum st}\right)\left(v/t - \frac{\sum stv/t}{\sum st}\right)\right] \\ &= r_{st:(u/s)\cdot(v/t)}\sigma_{st:(u/s)}\sigma_{st:(v/t)}\left(\sum st\right)^2, \end{aligned} \tag{5}$$

where $r_{st:(u/s)\cdot(v/t)}$ is the coefficient of correlation between the u_i/s_i and v_i/t_i with weights s_it_i and $\sigma_{st:(u/s)}$ and $\sigma_{st:(v/t)}$ are similarly weighted standard deviations. We may also express Δ in forms analogous to (5) containing the correlation coefficients $r_{uv:(s/u)\cdot(t/v)}$, $-r_{sv:(u/s)\cdot(t/v)}$, and

---

[2] W. F. Osgood, in his *Advanced Calculus* (1935), p. 306, follows W. Blaschke, *Vorlesungen über Differentialgeometrie* (1921), I, pp. 6 and 13, in attributing (4) to Lagrange. But Lagrange's contribution, included in a paper presented in 1773, is simply an algebraic identity corresponding to the case of (3) in which $n=3$, $s_i=t_i$, and $u_i=v_i$. Some writers (e.g., Courant and Goursat-Hedrick) ascribe only this case to Lagrange. See T. Muir, *Theory of Determinants in the Historical Order of Development* (2nd ed., 1906), I, pp. 33–37.

[3] For applications of weighted correlation coefficients, see the following articles by the writer in this JOURNAL: "The Difference between the Paasche and Laspeyres Index-Number Formulas," September 1941, pp. 343–350; "Further Notes on the Difference between Index-Number Formulas," December 1941, pp. 519–524; and "Index-Number Differences: Geometric Means," June 1942, pp. 271–274. See also the papers of L. von Bortkiewicz mentioned in the first of these articles.

$-r_{ut:(s/u)\cdot(v/t)}$. In each instance, Δ is numerically equal to a weighted product moment times the square of the sum of the weights. A useful relation is thus established between product moments and double sums.

To emphasize the symmetry between the correlation and double-sum forms of Δ, we may explicitly introduce into (3) the variables and weights entering into the correlation coefficient in (5):

$$\Delta = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} s_i t_i s_j t_j \left(\frac{u_j}{s_j} - \frac{u_i}{s_i}\right)\left(\frac{v_j}{t_j} - \frac{v_i}{t_i}\right). \tag{3a}$$

Similarly, we may rewrite (3) to correspond to any of the other weighted correlation coefficients appearing in alternative forms of (5). Incidentally, variants of (3) such as (3a) could, of course, be obtained directly from variants of (2) if the sums in $|\Delta|$ were first rewritten appropriately.

We now turn to an examination of some of the inequalities represented by Δ. First, we shall consider instances in which the sense of the inequality is not fixed.

An important example is the index-number inequality, which has already been discussed at some length in this JOURNAL.[4] If the u_i and s_i represent, say, alternative sets of quantity weights and the v_i and t_i represent prices in two periods, the latter being the base, then Δ gives the sign of the difference between two aggregative price indexes. The sign of the difference between the same two indexes converted to weighted arithmetic means of the relatives v_i/t_i is given by $\Delta = \sum ST \sum UV - \sum SV \sum UT$, where $S_i = s_i t_i$, $T_i = 1$, $U_i = s_i t_i (u_i/s_i)$, and $V_i = v_i/t_i$; the two sets of weights are U_i and S_i. This result is widely applicable, for any harmonic mean of relatives is also equivalent to a weighted arithmetic mean. In the special case where the u_i/s_i (i.e., the ratios of the alternative sets of weights for either the aggregative or mean forms) represent quantity relatives corresponding to the price relatives being averaged, we have the difference between the Paasche and Laspeyres indexes. From (3a), it is evident that the Paasche index is necessarily greater than the Laspeyres index if the coefficient of rank correlation between the two sets of relatives is $+1$, and necessarily smaller if the value of the coefficient is -1. Tchebycheff obtained an equivalent result for analogous integrals.[5]

The tetrad difference, $r_{a\cdot b} r_{c\cdot d} - r_{a\cdot d} r_{c\cdot b}$, is a fundamental notion of psychometrics and also occurs in a simple case of vector correlation.[6] It may be represented by Δ if the variables s_i, t_i, u_i, and v_i are replaced by $a_i' = (a_i - \sum a/n)/\sqrt{n}\sigma_a = (a_i - \sum a/n)/[\sum(a - \sum a/n)^2]^{1/2}$ and analogously defined b_i', c_i', and d_i'. The tetrad vanishes when a weighted correlation coefficient such as $r_{a'b':(c'/a')\cdot(d'/b')}$ or $r_{c'd'(a'/c')\cdot(b'/d')}$ equals zero; or, in geometric terms, when the planes a', c' and b', d' (the vectors have the coordinates a_i', c_i', b_i', and d_i', respectively) are at right angles.

---

[4] See papers by the writer mentioned in footnote 3.

[5] *Inequalities*, pp. 43–44.

[6] For the special case of vector correlation having the tetrad difference as numerator, see H. Hotelling, "Relations between Two Sets of Variates," *Biometrika*, December 1936, p. 324, or C. C. Craig,

The sign of the partial correlation coefficient $r_{ab\cdot c}$ is given by the sign of its numerator, $r_{a\cdot b} - r_{a\cdot c}r_{b\cdot c} = r_{c\cdot c}r_{a\cdot b} - r_{a\cdot c}r_{b\cdot c}$, which is in the form of Δ for the variables a_i', b_i', and $c_i' = d_i'$ defined as in the preceding paragraph (whence $\sum c'^2 = 1$). Thus, the sign of the partial coefficient corresponds to that of $r_{a'b':(c'/a')\cdot(c'/b')}$ or $r_{c'^2:(a'/c')\cdot(b'/c')}$.

The signs of the constants of the regression equation $Y = f + gX$ may also be considered to depend on expressions like Δ which form the numerators of the determinant solutions to the normal equations.[7] For example, if the observation equations are weighted by w_i, then the following substitutions in Δ give the numerator of the regression coefficient g: $s_i = w_i$, $t_i = 1$, $u_i = w_iX_i$, and $v_i = Y_i$. If the observation equations are unweighted, then $s_i = t_i = 1$, $u_i = X_i$, and $v_i = Y_i$. The corresponding weighted and unweighted correlation coefficients, $r_{w:X\cdot Y}$ and $r_{X\cdot Y}$, have the same signs as these Δ's.

We now turn to instances in which the two sides of (1) are "comparable"—in which the sense of the inequality is fixed for at least non-negative values of the variables. In each case, we shall make the left member of (1) equal to or greater than the right; that is, Δ will be positive except in cases of equality.

A familiar definitely directed inequality is the difference between the (weighted or unweighted) arithmetic and harmonic means of a set of positive quantities. Thus, if the z_i are positive, $\sum z \sum z^{-1} \geqq n^2$ is clearly valid, for the determinant products derivable from

$$\left\| \begin{matrix} \sum z & n \\ n & \sum z^{-1} \end{matrix} \right\| = \left\| \begin{matrix} z_1 \cdots z_n \\ 1 \cdots 1 \end{matrix} \right\| \cdot \left\| \begin{matrix} 1 & \cdots & 1 \\ z_1^{-1} & \cdots & z_n^{-1} \end{matrix} \right\|$$

are all non-negative.

Another common relation which may readily be verified by the methods described in this paper is the following moment inequality for (weighted) positive quantities:

$$\frac{\sum pz^m}{\sum p} \geqq \frac{\sum pz^k}{\sum p} \frac{\sum pz^{m-k}}{\sum p}, \qquad m > k > 0, \quad p_i > 0.$$

"Recent Advances in Mathematical Statistics, II," *Annals of Mathematical Statistics*, March 1942, p. 76. For vector treatment of tetrads, see E. B. Wilson, "On Hierarchical Correlation Systems," *Proceedings of the National Academy of Sciences*, 1928, pp. 283–291.

[7] The determinant in the denominator also has the form of Δ but may be shown by (3) to be necessarily positive; in fact, this determinant is an example of Cauchy's inequality, which is discussed several paragraphs later.

Decomposition of the matrix

$$\begin{Vmatrix} \sum p & \sum pz^k \\ \sum pz^{m-k} & \sum pz^m \end{Vmatrix}$$

shows that the typical determinant product is non-negative. This inequality may be extended by repeated application to the factors on the right-hand side.

The validity of the following moment inequalities, in which the weights p_i are positive, may also be established readily by examination of the typical determinant products:

$$\sum p \sum pz^{2m} \geqq (\sum pz^m)^2$$
$$\sum pz^{2k} \sum pz^{2m} \geqq (\sum pz^{m+k})^2$$
$$\sum pz^{m+1} \sum pz^{m-1} \geqq (\sum pz^m)^2.$$

The first and second of these inequalities hold for negative as well as positive values of z_i and of the exponents; the third holds for mixed values of z_i if m is an odd integer.

The most famous special case of (1) is Cauchy's inequality, which occurs not only in a variety of statistical problems but also in the derivation of more complex inequalities, like Hölder's and Liapounoff's. The matrix of Δ for Cauchy's inequality may be written as the product of two identical rectangular matrices; the typical determinant product is a perfect square. All but one of the definitely directed inequalities which have already been considered may be written in Cauchy's form. For example, the expression giving the sign of the difference between an arithmetic and a harmonic mean is equivalent to the following:

$$\sum z \sum z^{-1} - n^2 = \sum (z^{1/2})^2 \sum [(z^{-1})^{1/2}]^2 - [\sum z^{1/2}(z^{-1})^{1/2}]^2 \geqq 0.$$

Cauchy's inequality is equivalent to the statement that the cosine of the angle between two vectors is never greater than unity. Since any correlation coefficient may always be written in cosine form, its value obviously cannot exceed ±1, as is well known. By virtue of (3) or (3a), various correlation coefficients may be reduced to cosine form in more than one way; that is, the difference between the denominator and numerator may be given by more than one Cauchy inequality. For example, at least two such expressions exist for the weighted zero-order coefficient:

$$r_{w:a\cdot b} = \frac{\sum w(a - \sum wa/\sum w)(b - \sum wb/\sum w)}{[\sum w(a - \sum wa/\sum w)^2 \sum w\,(b - \sum wb/\sum w)^2]^{1/2}}$$
$$= \frac{\sum_i \sum_j w_i w_j (a_i - a_j)(b_i - b_j)}{[\sum_i \sum_j w_i w_j (a_i - a_j)^2 \sum_i \sum_j w_i w_j (b_i - b_j)^2]^{1/2}}.$$

In the unweighted case, we have

$$r_{a\cdot b} = \frac{\sum (a - \sum a/n)(b - \sum b/n)}{[\sum (a - \sum a/n)^2 \sum (b - \sum b/n)^2]^{1/2}}$$
$$= \frac{\sum_i \sum_j (a_i - a_j)(b_i - b_j)}{[\sum_i \sum_j (a_i - a_j)^2 \sum_i \sum_j (b_i - b_j)^2]^{1/2}}.$$

The first-order coefficient $r_{ab\cdot c}$ may be expressed in an alternative cosine form involving the a_i', b_i', and c_i' defined in an earlier paragraph:

$$r_{ab\cdot c} = \frac{r_{c\cdot c} r_{a\cdot b} - r_{a\cdot c} r_{b\cdot c}}{(r_{a\cdot a} r_{c\cdot c} - r_{a\cdot c} r_{a\cdot c})^{1/2} (r_{b\cdot b} r_{c\cdot c} - r_{b\cdot c} r_{b\cdot c})^{1/2}}$$
$$= \frac{\sum_i \sum_j (c_i' a_j' - c_j' a_i')(c_i' b_j' - c_j' b_i')}{[\sum_i \sum_j (c_i' a_j' - c_j' a_i')^2 \sum_i \sum_j (c_i' b_j' - c_j' b_i')^2]^{1/2}}.$$

Partial coefficients of higher order may be written in analogous fashion in terms of multiple sums of products of higher-order determinant pairs.

In conclusion, it may be well to recall at this point the analogy between the first-order correlation coefficient and the fundamental cosine law of spherical trigonometry.[8] Since the latter may be represented by a Lagrange identity,[9] so may the former. Thus, if $\mathbf{a}'$, $\mathbf{b}'$, and $\mathbf{c}'$ are vectors of unit length having the rectangular coordinates a_i', b_i', and c_i', respectively ($i = 1, \cdots, n; \sum a'^2 = \sum b'^2 = \sum c'^2 = 1$), then

$$(\mathbf{c}' \times \mathbf{a}') \cdot (\mathbf{c}' \times \mathbf{b}') = \mathbf{c}' \cdot \mathbf{c}' \mathbf{a}' \cdot \mathbf{b}' - \mathbf{a}' \cdot \mathbf{c}' \mathbf{b}' \cdot \mathbf{c}'.$$

The partial coefficient corresponds to the cosine of the dihedral angle between the planes $\mathbf{c}'$, $\mathbf{a}'$ and $\mathbf{c}'$, $\mathbf{b}'$. The scalar products $\mathbf{a}' \cdot \mathbf{b}'$, $\mathbf{a}' \cdot \mathbf{c}'$, and $\mathbf{b}' \cdot \mathbf{c}'$, which represent the cosines of the angles subtended by the sides of the spherical triangle, also correspond to the simple correlation coefficients $r_{a\cdot b}$, $r_{a\cdot c}$, and $r_{b\cdot c}$. The vector identity may readily be translated into the double-sum coordinate form of the partial coefficient shown in the preceding paragraph.

---

[8] See D. Jackson, "Trigonometry of Correlation," *American Mathematical Monthly*, June 1924, pp. 275–280, or R. W. Burgess, *Introduction to Mathematics of Statistics* (1927), pp. 275–276. For a comprehensive treatment of correlation in general, see J. MacMahon, "Hyperspherical Goniometry; and Its Application to Correlation Theory for *N* Variables," *Biometrika*, December 1923, pp. 173–208.

[9] See, for example, C. E. Weatherburn, *Elementary Vector Analysis* (1939), p. 72.

CONTENTS OF TIPPETT'S "RANDOM SAMPLING NUMBERS"

By Robert Gage, M.S.
Section on Biometry and Medical Statistics
Mayo Clinic, Rochester, Minnesota

As part of a plan for setting up a series of punch cards to be used for experiments with random samples, there were preliminarily punched onto Hollerith cards all the numbers in Tippett's tables of random sampling numbers.[1] In view of the wide use of this table for random experiments of various sorts, it was thought that it would be of some interest to report briefly a portion of the results obtained from tabulating these cards. Previous analysis of Tippett's data have been made by Yule[2] and by Dodd.[3]

The table contains 41,600 digits printed on twenty-six pages, each page containing thirty-two columns and fifty rows; on each page the columns are assembled in groups of four. The fifty digits of each column were punched onto a card, one card for each column, thus providing 832 cards that contained all the numbers in the table. In order to check the punching the entire procedure was repeated independently on another set of cards and checked mechanically.

Table I gives the frequency of each digit and a few summary findings for runs. The runs were counted mechanically on a tabulator as for one continuous series of digits constituted as follows. The top row of the first page was made continuous with the top row of the second page, this with the top row of the third page, and so forth, until the top row of the last page, 26, was completed. The latter was then made continuous with the second row of the first page, and the process continued through the fiftieth row of the last page to complete the series.

TESTS OF RANDOMNESS

The order of the digits can be examined, of course, by a great variety of statistical tests, each testing the "null hypothesis" that the digits are arranged at "random." The various tests that can be constructed will be differently sensitive to different kinds of systematizations,[4] and by noting which gives unlikely values for the test P, one might be able to discern what sort of systematization was present. A number of tests that were tried are given here as examples.

---

[1] L. H. C. Tippett, *Random Sampling Numbers*, Cambridge University Press, 1927.

[2] G. U. Yule, "A Test of Tippett's Random Sampling Numbers," *Journal of the Royal Statistical Society*, Vol. 101, pp. 167–172, 1938.

[3] E. L. Dodd, "Certain Tests for Randomness Applied to Data Grouped into Small Sets," *Econometrica*, Vol. 10, pp. 249–257, July–October 1942.

[4] Joseph Berkson, "Tests of Significance Considered as Evidence," this Journal, Vol. 37, pp. 325–335, September 1942.

J. Neyman, *Lectures and Conferences on Mathematical Statistics*, Washington, The Graduate School of the United States Department of Agriculture, 1938, Chap. 3, pp. 33–48.

1. The expected frequency for any digit and its standard deviation are given as pn and $\sqrt{pqn}$ respectively. Any observed frequency, for example, 4,264 for the digit zero, can be tested by the ratio (1) which will be normally distributed (for large n) in random samples with mean zero and unit standard deviation.

$$\frac{4{,}264-4{,}160}{\sqrt{0.1 \times 0.9 \times 41{,}600}} \quad 1.7, \qquad P = 0.09. \tag{1}$$

In this manner each of the digits can be tested separately to see whether there is any indication of preference or avoidance of particular digits.

2. An over-all test that will, however, not be so sensitive or revealing as to preference or avoidance of particular digits is the χ^2 frequency test which is given by (2).

$$\chi_q^2 = \frac{\sum (f - 4{,}160)^2}{4{,}160} = 11.2, \qquad P_{\chi_q^2} = 0.26. \tag{2}$$

3. *A priori* it would seem that tests based on "runs," that is, repetitions of the same digit, would be sensitive to likely departures from true randomization. It has been found by experience that it is practically impossible to maintain a random arrangement of "runs" in the presence of any kind of bias. An article by Mood[5] makes a number of such tests easily accessible.

For the number of runs of length j for digit i the mean and variance are given in (3) and the distribution for large numbers is sensibly normal.

$$\begin{aligned} E(r_{ij}) &= p_i{}^j q_i[(n - j - 1)q_i + 2] \\ \sigma_{r_{ij}}{}^2 &= p_i{}^{2j} q_i{}^2\{(n - 2j)(n - 2j - 1)q_i{}^2 + (n - 2j)q_i(1 + 5p_i) \\ &\quad + 6p_i{}^2 - [(n - j - 1)q_i + 2]^2\} \\ &\quad + p_i{}^j q_i[(n - j - 1)q_i + 2]. \end{aligned} \tag{3}$$

For instance, to test the runs of two of digit zero (Table I)

$$E(r_{02}) = (0.1)^2(0.9)[(41{,}600 - 2 - 1)(0.9) + 2] = 337.0,$$

$$\sigma_{r_{02}}{}^2 = 321.2$$

$$\frac{362 - 337}{\sqrt{321.2}} = 1.4, \qquad P = 0.08.$$

---

[5] A. M. Mood, "The Distribution Theory of Runs," *Annals of Mathematical Statistics*, Vol. 11, pp. 367-392, December 1940.

That is, the probability of getting 362 runs or more of two in random samples is 0.08. If there were an abnormal tendency to pairing of this digit, one would be likely to get a much smaller value for P using this test. Hence, there is little reason to suspect any abnormal pairing here. Similarly, each of the observed frequency of runs for the different digits can be tested.

TABLE I

FREQUENCY AND "RUNS" OF EACH DIGIT

i Digit	Frequency	r_{ij}, runs of length j					r_i Total runs
		1	2	3	4	5	
0	4,264	3,406	362	39	3	1	3,811
1	4,081	3,248	356	35	4	0	3,643
2	4,233	3,452	333	29	7	0	3,821
3	4,224	3,408	344	37	3	1	3,793
4	4,171	3,377	330	38	5	0	3.750
5	4,144	3,356	329	42	1	0	3,728
6	4,051	3,262	339	33	3	0	3,637
7	4,084	3,311	325	33	6	0	3,675
8	4,144	3,328	366	24	3	0	3,721
9	4,204	3,409	332	37	5	0	3,783
Mean	4,160	3,355.7	341.6	34.7	4.0	0.2	3.736.2
Expected mean	4,160	3,369.6	337.0	33.7	3.4	0.3	3,744.0

4. An over-all test for the runs of any particular digit can be made from the total runs observed for the digit. The expected value and variance are

$$\begin{aligned} E(r_i) &= np_iq_i + p_i^2 \\ \sigma_{r_i}^2 &= np_i(1 - 4p_i + 6p_i^2 - 3p_i^3) + p_i^2(3 - 8p_i + 5p_i^2). \end{aligned} \tag{4}$$

For example, to test the total runs of digit zero (Table I)

$$E(r_0) = 3{,}744, \qquad \sigma_{r_0}^2 = 2{,}733$$

$$\frac{3{,}811 - 3{,}744}{\sqrt{2{,}733}} = 1.3, \qquad P = 0.10.$$

Similarly, each of the other digits can be separately tested.

5. An over-all test can be made for the total runs for all the digits. It has been shown that the variable[6]

$$Q = \sum_1^k \sigma^{ij}x_ix_j \tag{5}$$

is asymptotically distributed according to the χ^2 law with k degrees of freedom. In (5) the x_i are defined as

---

[6] Formula 5 is a quadratic form and the j is not the length of run as in 3.

$$x_i = \frac{r_i - np_iq_i}{\sqrt{n}}, \qquad i = 1, 2, \cdots, k$$

and $\|\sigma^{ij}\|$ is the inverse of

$$\begin{aligned} \sigma_{ii} &= p_iq_i - 3p_i^2q_i^2 \\ \sigma_{ij} &= -p_ip_j(1 - 2p_i - 2p_j + 3p_ip_j). \end{aligned}$$

The x_i are readily determined from the data given in Table I, with $n=41{,}600$; $p_i=0.1$; $q_i=0.9$ and $k=10$. For this particular set of data $\|\sigma^{ij}\|$ becomes

$$\begin{aligned} \sigma^{ii} &= (2 - 3p)/p(1 - p)(1 - 2p) \\ \sigma^{ij} &= (1 - 3p)/p(1 - p)(1 - 2p), \qquad i \neq j. \end{aligned}$$

Substituting the numeric equivalents in (5), $Q=15.3$ which as a value of χ^2 for ten degrees of freedom gives $P=0.12$.

6. A quick over-all test can be made for the runs of all the digits in terms of the expectation and variance for such runs. The values are

$$\begin{aligned} E(r) &= n(1 - \sum p_i^2) \\ \sigma_r^2 &= n[\sum p_i^2 + 2\sum p_i^3 - 3(\sum p_i^2)^2]. \end{aligned} \tag{6}$$

Thus, the test for the 37,362 runs in this series is

$$E(r) = 37{,}440, \qquad \sigma_r^2 = 3{,}744$$

$$\frac{37{,}362 - 37{,}440}{\sqrt{3{,}744}} = -1.3, \qquad P = 0.10.$$

The fact that a variety of tests have not disclosed any indication of systematization in runs does not mean that none exists. It may be that the particular tests used were not sensitive to the particular type of systematization which existed. The kind of systematization which would be introduced is likely to be related to the manner in which the table was constructed. Since the digits are grouped in the table in groups of four, the possibility that there is systematization within these groups suggests itself. Table II gives the number of runs for each digit found in the 10,400 groups of four digits in Tippett's table. The runs may be tested by the same theorems as utilized before or by simple extension of them.

TABLE II

OBSERVATIONS OF "RUNS" IN 10,400 GROUPS OF FOUR DIGITS

	Mean		Digit									
	Expected	Observed	0	1	2	3	4	5	6	7	8	9
Runs of one	3,556.8	3,547.5	3,582	3,431	3,666	3,608	3,564	3,540	3,474	3,477	3,534	3,599
Runs of two	271.4	274.7	303	291	256	276	266	266	259	273	282	275
Runs of three	18.7	19.7	24	20	17	20	25	24	17	19	14	17
Runs of four	1.0	1.0	1	2	1	1	0	0	2	1	1	1

7. As an example, to test the number of runs of two of digit zero (Table II) the expected value for a single group of four digits and the variance are given by (3)

$$E(r_{02}) = 0.0261, \qquad \sigma_{r_{02}}^2 = 0.0254.$$

For the sum of such runs in 10,400 groups the expected value and variance are

$$E(r) = 0.0261 \times 10{,}400 = 271.4$$

$$\sigma^2(r) = 0.0254 \times 10{,}400 = 264.2$$

$$\frac{303 - 271.4}{\sqrt{264.2}} = 1.9, \qquad P = 0.03.$$

The observed mean value for runs of two for all ten digits was 274.7 as compared with an expected value of 271.4. A test[7] gives

$$\frac{274.7 - 271.4}{\sqrt{\dfrac{271.4}{10}}} = 0.6, \qquad P = 0.27.$$

Thus there is no evidence here either of systematization. Other tests have been applied with similar results.

I wish to express my thanks to Dr. Joseph Berkson for suggesting this survey and for guidance in its conduct.

---

[7] The correlation existing between the variables has been disregarded.

QUALITY CONTROL APPLIED TO BUSINESS ADMINISTRATION*

By Wm. B. Rice
Plomb Tool Company, Los Angeles

The techniques of quality control, as established by Shewhart,[1] Simon,[2] and others, are being used extensively and with great success in the manufacturing problems of industry, but it is not widely recognized that these techniques can be applied to the problems of office and financial management (cf., however, footnote 7).

It has well been said that quality control is one-tenth statistics and nine-tenths engineering judgment. If, in considering the non-manufacturing problems of business administration, executive judgment is substituted for engineering, it is possible to set up a criterion by which the usefulness of the application of quality control methods can be gauged. That criterion should require the purely statistical techniques to reinforce the seasoned conclusions of experienced, expert business managers. This test is being met successfully in several experiments conducted by the author, at the same time uncovering significant facts which executives can use in the direction of efficiency and profit.

Under the pressure of war-time conditions, many firms with contracts negotiated at fixed prices are finding that overtime wages paid to the office force become a noticeable factor in rising costs, a factor for which it is difficult to get compensation through the process of re-negotiation. The *preliminary* data used by the author in establishing economical limits for office overtime pay consisted in the percentage of pay for overtime work to pay for regular work, by weeks, for six departments, from January through June 1942.

If we think of each department as a different machine operator, and take for each "production sample" the figures for the four (or five) weeks in each month, we get six points for each of the six months, a total of thirty-six samples in all. The preliminary data were set up in tabular form as shown in the table. Beginning with July, the data were improved by eliminating a non-homogeneous element—the salaries of supervisory personnel whose compensation was on a monthly basis and who did not get overtime pay. Calculations made from the new figures

---

* A report presented at an Engineering, Science and Management War Training Course on Quality Control held at Los Angeles in the Fall of 1942 under the auspices of Stanford University, and sponsored by the U. S. Office of Education. The original presentation has been revised to carry it through December 1942.

Credit and thanks are due to Professor Holbrook Working, of Stanford University, and to Dr. W. Edwards Deming for their inspiring help in preparing this paper for publication.

[1] W. A. Shewhart, *Economic Control of Quality of Manufactured Product*, D. Van Nostrand, 1931.

[2] Leslie E. Simon, *An Engineers' Manual of Statistical Methods*, John Wiley & Sons, 1941.

for July, August, and September, raised the overall average from 13.4 to 20.3 per cent, reflecting a truer picture of the facts; and decreased the average range, as was to be expected, from 12.4 to 10.6 per cent.

RELATION OF OVERTIME PAY TO PAY FOR REGULAR WORK, BY WEEKS AND DEPARTMENTS

(*Per cent*)

	Weeks						
	1st	2nd	3rd	4th	5th	Average ($\bar{X}$)	Range (R)
January 1942							
Department 1	20.4	16.6	11.4	12.2	—	15.2	9.0
2	28.5	2.4	3.7	7.4	—	10.5	26.1
3	25.0	11.4	10.4	9.7	—	14.1	15.3
4	9.5	16.0	8.2	11.2	—	11.2	7.8
5	22.3	27.0	24.5	36.0	—	27.4	13.7
6	7.5	10.4	7.9	12.4	—	9.6	4.9
February 1942							
Department 1	17.1	16.1	7.8	9.9	—	12.7	9.3
2	5.1	5.6	4.2	7.8	—	5.7	3.6
3	11.6	10.2	7.6	16.3	—	11.4	8.7
4	12.2	15.5	11.6	12.0	—	12.8	3.9
5	21.2	18.3	34.3	48.3	—	30.5	30.0
6	15.9	15.2	11.3	15.1	—	14.4	4.6
			Etc.				

Using the new data for analyzing past experience, and the notation of the American Standards Association 1942 pamphlet[3] the central lines and action limits were computed as shown below.[4,5]

$$\bar{\bar{X}} = 20.3 \text{ per cent}$$

$$\bar{R} = 10.6 \text{ per cent for samples of 4}$$

$$\bar{\bar{X}} \pm A_2\bar{R} = 28.0 \text{ per cent and } 12.6 \text{ per cent}$$

$$D_3\bar{R} = 0$$

$$D_4\bar{R} = 24.2 \text{ per cent}$$

For several months in the first half of the year department 5 was "out of control" on both average and range (see chart). In June, an alert office manager caught the excessive overtime of department 5, hired more help, and promptly thereafter brought it back within normal limits. Needless to say, if the control chart technique had been available, indications of trouble would have been given much earlier, with a worth-while saving in labor costs.

---

[3] American Standards Association, *Control Chart Method of Controlling Quality During Production*, Z1.3-1942 (American Standards Association, 29 West 39 Street, New York), p. 39.

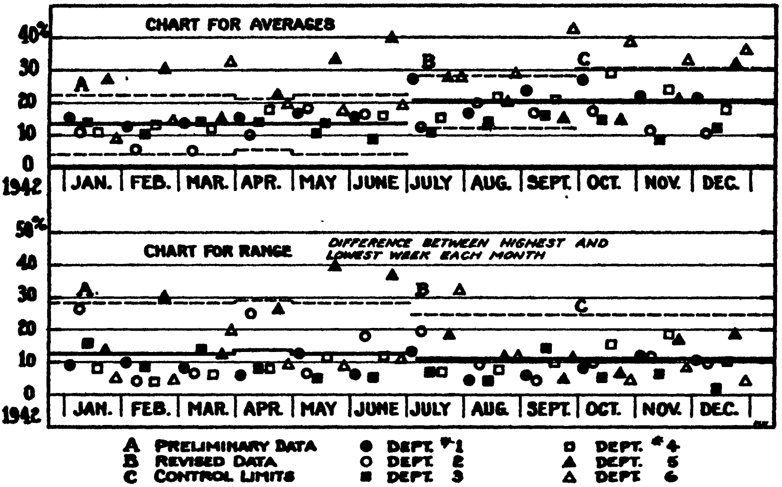
[4] The weeks containing July 4th and Labor Day were omitted because they were abnormal. Consequently, all three months contained four weeks.

[5] The quantities A_2, D_3, and D_4 are needed for computing the control limits on the charts. Their numerical values are given in tables; e.g. in Simon, *loc. cit.*, pp. 204 ff., and in the appendix to the American Standards publication listed in footnote 3.

More recently department 6 has gone out of control; but the difficulty of finding enough trained men for a specialized type of work under present labor market conditions is hindering efforts to remedy the situation. This, however, is not so much a statistical as a personnel problem.

CONTROL CHART FOR OFFICE OVERTIME
PER CENT OF OVERTIME PAY TO REGULAR PAY

(Each point is the average for one month)



Management's opinion as to the most economical amount of overtime allowable varies according to circumstances. Normally, the standard might be set at no overtime, i.e., $\overline{\overline{X}}=0$. Under war conditions, when volume of production is paramount and workers are hard to get, a certain amount of overtime work is often necessary. But the decision is difficult as to what limits should be set. They should certainly not be determined without operating facts to go on, just as engineering specifications for a manufactured article should not be drawn up without a knowledge of what limits inhere in the process.

In both factory and office quality control methods develop a basis upon which to base executive decisions.

How well such an application to overtime wages meets a common-sense criterion is indicated by the upper limits for averages derived from the data for July, August, and September. Now it happens that this limit, 28.0 per cent ($\bar{X}+A_2\bar{R}$ above), is close to 30 per cent, which represents 20 per cent overtime in hours worked, or one full extra day in a normal five-day work week. In this case the management considered it reasonable to argue that if any department had enough work to keep its staff busy for an extra day, economy would be served by hiring sufficient additional help to do the work in five days, thus avoiding the necessity of paying overtime premiums regularly.[6] With these thoughts in mind, the control limits for future data were set up as follows:

For averages:	Upper limit, 30 per cent
	Lower limit, 0 per cent
For ranges:	Upper limit, 24.2 per cent
	Lower limit, 0 per cent

These control limits were applied to the last quarter of the year, i.e., October to December 1942. The weeks containing holidays were omitted from the calculations because, with a full working force present on those days, overtime payments were necessarily abnormal. Effectiveness of quality control is indicated by the fact that with the exception of department 6 and one point for department 5, all overtime was kept within the agreed-upon limits. Seasonal factors accounted for department 5's unavoidable lapse in December.

In this connection the range chart is particularly important. Showing as it does the difference between the highest and lowest week in amount of overtime paid for during the month, the range gives a measure of the skill and care with which work-flow through each department is planned and supervised. Extreme irregularity in work-flow shows up as an out-of-control point on the range chart, and can often be used to disprove a department head's alibi of "too much work" when he is confronted with the fact of excessive overtime paid to his employees.

The author's observation is that, in applying quality control to general management, the behavior of the range or standard deviation chart will reveal more valuable information than the chart for averages.

Office overtime analysis is only one example of the application of quality control to the interests of general management. The field is fertile but largely unexplored. For instance, quality control can be used to set norms and action limits for clerical work, both by individuals and by groups where, as in stenography or billing, quantitative standards can be established. An application toward the improvement of clerical work, and toward reduction in costs, has been made and reported by Dr. Deming and his staff at the Bureau of the Census.[7] Evaluation of sales and territories, by salesmen and by lines, can be undertaken. Budgetary and cost accounting records are illuminated. Along some of these lines work now being done by the author indicates that, as in manufacturing, such studies do determine economic action limits, thereby lessening the heavy burdens of executive responsibility.

---

[6] With the 48-hour, six-day week established officially in many sections of the country, these conclusions may need to be reconsidered.

To the many friends of quality control who are personally convinced of its value but who have to sell their superiors on the idea, its application to specific problems in business administration offers an inexpensive and powerful psychological approach. It requires no costly experimental interruptions of production; treads on few toes; invites fewer organizational upheavals; and by proving its worth in one field, opens the way logically to many others.

---

[7] W. Edwards Deming and Leon Geoffrey, "On Sample Inspection in the Processing of Census Returns," this JOURNAL, Vol. 36, 1941, pp. 351–360.

FAMILY VERSUS INDIVIDUAL IN MEASUREMENT OF AUDIENCES

BY ALFRED POLITZ

ADVERTISING RESEARCH owes its place entirely to the requirements of advertising itself. Thus the value of a particular research tool depends entirely and only on its usefulness to advertisers.

Since the best of advertising has no selling power until it is presented, one of the most fundamental problems to advertisers, publishers and radio men is to know how many people read a magazine or listen to a radio program. And it is in solving this basic problem that a number of research endeavors have adopted a method which defeats their purpose.

It came about this way. Since most people live together in groups as families, there exists the idea of the family as a social unit. But this idea goes a step further as far as advertising is concerned. Many or most articles which the members of the family consume are not bought by all members separately, but only by one representative, for example, the housewife. Thus the family assumes the appearance not only of a social unit, but also of a single consuming unit.

In view of this, it seemed to research men and publishers, and justly so, that it was inadequate to express the readership of an advertising medium in terms of total individuals. It seemed that there should be a unit of measurement which would better parallel this structure of consumption. So the "family base" was chosen. No specific studies using the family as the base need to be mentioned, since it is generally known that many, if not most, readership reports employ this method.

Assume that there is a statistical universe of 24 families. For the sake of simplicity, each family consists of four members: Husband, housewife, son, and daughter.

In this universe there are two newspapers, A and B. Expressed in terms of individuals, each newspaper has 48 readers. So in terms of individuals they are equal.

But the distribution of the individuals differs in the following way:

A is read by 2 persons in each of
24 different families (housewife and daughter);

B is read by 4 persons in each of
12 different families.

Which paper is a better medium for tomato juice?

According to certain research conclusions, A is better, because it goes into twice as many families as B.

The idea is that tomato juice is purchased by the housewife. If the ad is read by her, the purpose of the ad is practically fulfilled. Any additional reading by husband, daughter, or son is of only secondary value. These persons may speak to the housewife about tomato juice and increase her willingness to buy a certain brand, but it is obvious that such influence upon the final purchase is smaller than that of the housewife herself.

A and B are equal, inasmuch as each is read by 48 persons.

Newspaper A, however, which goes into 24 families, influences 24 housewives. While newspaper B, which goes into 12 families, influences only 12 housewives. And so, on the "family base," A is better than B.

Thus there are many research men and advertisers who believe that the family, instead of the individual, is the right unit for describing the readership of newspapers and magazines and the audience of radio programs. But the whole family approach is based on a fallacy.

It is true that, in this particular example, the number of families comes closer to the description of the relative values of the newspapers than does the total number of individuals. But that is more or less accidental and does not warrant the use of "the family" as a yardstick. *For what really was used as the gauge was the number of housewives.*

The family is only a condition of social life which gives the housewife (an individual) a greater importance for buying tomato juice than the son (another individual). Consequently an advertiser of tomato juice is more interested in the housewife than in the son.

There are other conditions, also, which determine the value of certain individuals.

Some people's habit of living in apartment houses makes the superintendents, who also live there, more important to coal dealers than the ordinary tenants. Southerners live in a warmer climate than the Northerners, and so are more important to manufacturers of light clothing.

There is a direct and simple device in every statistical picture which takes care of situations in which one type of individual is distinguished from other types. That is the breakdown of the total figure into specific ones.

Suppose, for instance, that the total figures of readers are known for different magazines, and that Southerners are more important for the given purpose than Easterners, Westerners, and Northerners. Then, from the total figures, the number of Southerners within the readership of each magazine simply must be sorted out.

There is a universe of people living in apartment houses. Two newspapers are distributed among them, and both newspapers are equal in the total number of individual readers. But the particular advertising problem is concerned with coal. Therefore the total number of readers is not sufficient. The figure must be broken down into "superintendents" and "tenants." If it appears that one paper is read by more superintendents than the other, it therefore has a better chance to sell coal.

If lipstick is the product, the readership figures are broken down into females and males. For Cadillac cars, they are broken down into well-to-do people and not well-to-do.

For tomato juice, the figures must be broken down into housewives and others.

If paper A with 48 readers reaches 24 housewives, and B with its 48 readers reaches only 12 housewives, then A is better than B. It is better because it reaches more housewives and not because it reaches more families.

It is possible that the greater number of housewives may coincide with a greater number of families. And in the first "tomato juice" example, it actually did.

Any different example, however, is equally justified.

Suppose, as in the first example, that paper A has 48 readers in 24 different families. And it is always read by two family members—in this case, however, son and daughter, instead of housewife and daughter.

Paper B is read entirely as before. It has 48 readers in 12 families, and it is always read by all four family members, husband, housewife, son, and daughter.

If, for the sake of tomato juice, the number of *families* indicates the value of a newspaper, then A is still better than B, for A still reaches twice as many families.

But looking at the kind of *individuals* it reaches—there is not a single housewife involved. Although there are twice as many families, it is read only by sons and daughters. So it cannot have twice as great a value; quite on the contrary, it may have no value at all.

Therefore as was said before, if newspaper A has more value for tomato juice, it cannot be on account of the greater number of *families*, but only on account of the greater number of *housewives* it reaches. And it is a mere accident if both figures happen to be in agreement. (As a matter of fact, they usually are in disagreement, for the very reason that a family is a collective of heterogeneous elements in which all ages and sexes are mixed.)

The family, itself, has neither a sex nor an age nor an occupation in which an advertiser might be interested. It cannot read and it cannot buy, the two activities which are most important to an advertiser. And whenever an attempt is made to use such a unit as a yardstick in measuring activities which it does not perform, but which are performed rather by one or two or three or more of its members, the purpose of the yardstick is defeated.

If surveys using the family base speak of 1000 families which are reached by a newspaper or magazine, they do not define what they mean. Do two members of each family read the newspaper? Or does one member of each family read the paper? Or is a family counted if at least one member reads the paper? In practice the "at least one" definition is used.

"At least one" means: One, or two, or three, or more. There is the first variable; it is not known how many people read the newspaper.

But there is still another variable. If a paper reaches 1000 families—and that means it reaches either 1000 or 3000 or more people—who are these people? They are not housewives, for if they were, they certainly would have been called so. Nor, for the same reason, are they husbands. They are, in fact, not described at all.

In a short sentence it looks like this: A newspaper which "reaches 1000 families" is a newspaper which is read by 1000 or 3000 or more people, who are either housewives or husbands or someone else.

The variables in this definition render the "family" useless as a yardstick. There is a possibility, however, of making the "family" useful as a yardstick by attaching to it a sufficiently rigid definition. That would be: A family is reached by a newspaper or radio program if each of its members (or adults) is reached.

On the basis of this reasonable definition, however, what must the research man do with a copy which is not read by each member of a family, but by only one or two or three members of a family which consists of four? In accordance with the last definition, such a copy does not reach the family.

The number of families reached would be correct and clear in itself, but very small. Only a minority of copies reaches every member of a family. So by this definition the majority of copies is neglected. In other words, the figure so obtained does not say anything about the fate of the copies, but only about the fate of some of the copies. It does not tell about readers of the paper; it tells only about certain readers, namely, those who have close relatives reading the same paper. The definition is technically clear but irrelevant to the purpose. This is so obvious that it automatically explains why this definition never was used.

It must be remembered that there is a priori so much clarity and simplicity in expressing the value of a newspaper or magazine in terms of individuals, that the inventors of the family base needed an enticing reason for becoming involved in it. The enticing reason was the fact that equal numbers of individual readers do not necessarily express equal values of newspapers. Paper A reaches 48 readers in 24 families (husband and housewife are readers). Paper B reaches 48 readers in 12 families (husband, wife, son, and daughter are readers). For selling tomato juice, newspaper A is better than B. "While having the same total of individual readers, it reaches twice as many families." Now this statement is only true if a family can be constituted by two of its members as well as by four.

Thus it develops that the "family" could offer its delusive service as a yardstick only *because* of its hidden flexibility.

It is understood that no one tried to aim at flexibility. The aim was to use a unit which would parallel in audience measurement the known facts of family consumption. Unfortunately "the family" was chosen as the unit, without sufficient analysis. It provided the chance to use comfortable terminology and gave the feeling of scientific rigor. Yet by its very nature it led to erroneous conclusions.

On the other hand it requires only a very small sacrifice of preliminary comfort to obtain and enjoy an accurate method of measuring reading and listening audiences. It is the individual who listens and reads; it is a certain type of individual whose reading and listening are more or less important for a particular purpose. And it is the statistical sorting out of these types which tells the complete story about the value of a radio program or a periodical.

THE COST OF LIVING IN CITIES IN RELATION TO THEIR SIZE AND LATITUDE

By E. L. Thorndike and Ella Woodyard

IT IS REPEATEDLY ASSERTED or assumed that the cost of living is much greater in large cities than in small, and in northern cities than in southern. Both statements are false. Except for rent the differences in retail prices between cities of 30,000 to 44,000 inhabitants and cities of 300,000 to 500,000 inhabitants are near zero and the difference in rent for a house or apartment of equal quality is less than twenty per cent. The differences in relation to latitude are almost nil.

We examined the advertisements appearing September 13 to 18, 1937, in the daily newspapers of the cities (265 in number) of over 30,000 population and under 500,000 (in 1930) which had daily newspapers.[1] We recorded carefully the prices found in all advertisements for any of the 28 items listed below, except that care was taken not to include more than once an apartment or house advertised on more than one day or by more than one agent, and that if more than twenty of any one size were advertised, twenty were chosen at random.

1. Monthly rent, apartment of 3 rooms, unfurnished
2. Monthly rent, apartment of 4 rooms, unfurnished
3. Monthly rent, apartment of 5 rooms, unfurnished
4. Monthly rent, apartment of 6 rooms, unfurnished
5. Monthly rent, house of 4 rooms, unfurnished
6. Monthly rent, house of 5 rooms, unfurnished
7. Monthly rent, house of 6 rooms, unfurnished
8. Monthly rent, house of 7 rooms, unfurnished
9. Monthly rent, house of 8 rooms, unfurnished
10. Sale price of house of 4 rooms, unfurnished
11. Sale price of house of 5 rooms, unfurnished
12. Sale price of house of 6 rooms, unfurnished
13. Sale price of house of 7 rooms, unfurnished
14. Sale price of house of 8 rooms, unfurnished
15. Sale price of house of 9 rooms, unfurnished
16. Permanent wave, lowest of the advertised prices
17. Permanent wave, median of the advertised prices
18. Permanent wave, highest of the advertised prices
19. Movies, average of lowest price of evening admission, and of other than evening admission if such is advertised
20. Movies, average of median price of evening admission, and of other than evening admission if such is advertised
21. Milk, using all firms advertising[2]

---

[1] Except that three resort cities (Atlantic City, Miami, and St. Petersburg) were excluded and Milwaukee was used instead of Washington.

[2] Milk prices appeared in too few cities (only 18) to be used.

22. Butter, using all firms advertising
23. Eggs, using all firms advertising
24. Chicken, using all firms advertising
25. Fowl, using all firms advertising
26. String beans, using all firms advertising
27. Potatoes, using all firms advertising
28. Cabbage, using all firms advertising

Expressing all prices as percentages of the median price for the thing in question in cities of 55,000 to 65,000, we have the facts of Table I.

TABLE I

THE RELATION OF CERTAIN ADVERTISED PRICES TO SIZE OF CITY IN 265 CITIES

Size of city	Prices expressed as percentages of the median price for cities of 55,000 to 64,999 inhabitants							Average
	Chicken	Fowl	Vegetables	Eggs	Butter	Movies	Permanent waves	
300,000 to 499,999	85½	94	100	103	104	100	106	98.9
200,000 to 299,999	94	100	110½	89	103	102	90	98.4
115,000 to 199,999	104	100	100	96	103	101½	99	100.5
95,000 to 114,999	96	104	100	98	100	91½	100	98.5
75,000 to 94,999	96½	93	105½	104	100	100	100	99.9
55,000 to 74,999	100	100	107	98	100	100	99½	100.6
45,000 to 54,999	100	100	107	96	101	103	107	102.0
35,000 to 44,999	100	100	107	104	100	85	102	99.7
30,000 to 34,999	94	96	100	96	100	100	97	97.6

We did not make similar records and computations for the retail prices of staples like flour, sugar, gasoline, trade-marked goods, drugs, books, cigars or cigarettes, because these should be less sensitive to size of city than the commodities of Table I. But a rough inspection of prices for some of them shows no evidence that they are cheaper in the smaller cities.

Expressing the median rental of an apartment or house of specified size in each group of cities as a percentage of the median price in cities of 55,000 to 65,000, we have nine sets of percentages for apartments of 3, 4, 5, and 6 rooms and houses of 4, 5, 6, 7, and 8 rooms. We also have six similar sets of percentages for sales prices of houses of 4, 5, 6, 7, 8, and 9 rooms. The averages, when all fifteen sets are given equal weight, are as shown in Table II. There is a clear rise of about 10 per cent from cities under 55,000 to cities of 55,000 to 114,999, and a less clear rise of about 5 per cent from cities of 55,000 to 114,999 to cities of 115,000 or over. Any reasonable system of weights gives closely similar results. In so far as the apartment or house of the same number of rooms tends to be of better quality in the larger cities, the figures of Table II exaggerate the differences in rental or purchase for equally good accommodations.

TABLE II

THE RELATION OF ADVERTISED PRICES OF APARTMENTS AND HOUSES FOR RENT AND FOR SALE TO SIZE OF CITY IN 265 CITIES

Size of city	Average of fifteen sets of percentages for rents of apartments of 3, 4, 5, and 6 rooms, and houses of 4, 5, 6, 7, and 8 rooms, and of houses for sale of 4, 5, 6, 7, 8, and 9 rooms
300,000 to 499,999	102½
200,000 to 299,999	98
115,000 to 199,999	98
95,000 to 114,999	91
75,000 to 94,999	95
55,000 to 74,999	98½
45,000 to 54,999	86
35,000 to 44,999	86
30,000 to 34,999	84½

The higher cost of lodging in the larger cities is not a consequence of higher prices for building materials. The data reported by the U. S. Bureau of Standards on the prices of brick, sand, stone, cement, lime, boards, 2×4's, gypsum board, nails, tar-paper, glass, linseed oil and white lead, are in fact higher for cities under 100,000 population than for those from 100,000 to 200,000, and higher for the latter than for cities of over 200,000. Using percentages of a base-price for each commodity, the averages of the thirteen percentages are as shown in Table III.

TABLE III

PRICES OF BUILDING MATERIALS IN RELATION TO SIZE OF CITY

Size of city	Number of cities studied by the Bureau of Standards	Average of 13 percentages for brick, sand, etc.
Over 1,000,000	4	92.9
500,000 to 999,999	6	91.7
300,000 to 499,999	8	92.3
200,000 to 299,999	9	96.5
100,000 to 199,999	16	100.2
30,000 to 99,999	11	102.6
Under 30,000	3	114.9

A critical study of the costs of lodging, foods, carfares, electric current and other commodities and services in various cities reported by the U. S. Bureau of Labor Statistics and the National Industrial Conference Board reveals that the discrepancies between their findings concerning size of city and ours are not large and may be explainable by differences in quality.

The advertised prices in the 265 cities show no clear differences in relation to latitude, either for foods, movies, etc., or for rent. The medians of eleven percentages summarizing both for the groups at 44° or more, 43°, 42°, 41°, 40°, 39°, 38°, 36° and 37°, 33° to 35°, and south of 33° are, in order, 96, 91, 100, 100, 102, 100, 98, 102, 93, and 88.

In the prices for domestic use of electricity reported for 51 cities by the Bureau of Labor Statistics there is little or no correspondence with latitude. The percentages for 12 cities north of 42°, 13 cities at 40° or 41°, 13 at 37° to 39°, and 13 south of 37°, are in order 102, 104, 90, and 105. The prices of building materials are a little *lower* in the most northerly cities.

EXPERIENCE WITH TESTS OF SIGNIFICANCE: A REPLY TO PROFESSOR R. A. FISHER

BY JOSEPH BERKSON, M.D.
*Section of Biometry and Medical Statistics, Mayo Clinic**

PROFESSOR R. A. FISHER in his note on my article "Tests of Significance Considered as Evidence"[1] finds it pertinent to say, "It is not my purpose to make Dr. Berkson seem ridiculous, nor, of course, to prevent him from providing innocent amusement," and to make some biographical comments of a similar intention. This is no time to begin calling names across the sea, and however strongly I may differ with Professor Fisher in regard to scientific questions, I shall confine my differences to the subject matter discussed.

My present attitude toward statistical tests of significance is the result of more than fifteen years of experience in their application to a large variety of data. I have no desire to undermine them unwarrantably. On the contrary, since my profession is that of consulting statistician, I should like to possess a tool in the use of which I was a specialist and which could serve well those who engaged my services. But if I have found these tests misleading in certain applications, it is appropriate that I expose this fact in statistical journals for the view of my colleagues. I may be wrong, but I am not such a fool that I cannot repeat as well as anyone who has ever memorized a textbook, the argument about the null hypothesis and the small P disproving it. At the beginning of my career I quoted the orthodox view. But a consulting statistician in a practical institution finds himself under the necessity of reconciling what *he says* to those who employ him with what *they find*. Time and time again I came upon actual situations in which (1) it was patent to all competent people who viewed the data, that the null hypothesis, on the evidence, *was not* true, but a small P did not result, (2) a small P did result, but it was obvious to all who examined the data, that on the evidence the null hypothesis *was* substantially true, or (3) when the null hypothesis was tested by several tests, each put forward with equally good authority, the P's were considerably *different*.

Reflection on this situation led me to the conclusion that the argument by which tests of significance are usually supported is basically unsound. The fallacy lies fundamentally in the argument's not considering the *alternative* to the null hypothesis. I have found that the quite numerous instances of false or paradoxical conclusions of my experience are resolved, if the view is adopted that what the result obtained from a test of significance does is not to "disprove the null hypothesis tested," nor even to give a basis for "rejecting" it but rather that the finding affords evidence *in favor* of something—in favor of some *alternative*. In ordinary rational procedures and in scientific inquiry too, one posits explicitly or implicitly a set of alternatives, A and not A: for instance, the patient has syphilis or he does not have syphilis; there is no difference in means or there is some difference; the distribution is mesokurtic or the distribution is heterokurtic; the regression is straight or the regression is curved. When the P of an appropriate t test is small, according to my view, one does not "disprove" that there is no difference in means; one has affirmative evidence in favor of the alternative, that is, that there is such a difference. I say this constitutes such evidence, because in the circumstances, a small P would occur frequently (relative to the level of significance) if there did exist a difference between the means.

---

* At present on leave, Lt. Col., Medical Corps, Army of the United States, Chief, Statistical Division, Office of the Air Surgeon.

[1] Joseph Berkson, "Tests of Significance Considered as Evidence," this JOURNAL 37: 325–335 (September) 1942. R. A. Fisher, "Note on Dr. Berkson's Criticism of Tests of Significance," this JOURNAL, 38: 103–104 (March) 1943.

The small P's of tests are evidence for any alternative in which they occur frequently. Good tests are ones in which (1) when a small P occurs it is evidence for an alternative in one particular pair of contraries and (2) when this alternative is true, a small P will occur very frequently. Such a test might be called highly specific and sensitive. An example is the t test appropriately performed with ample numbers for a difference between the means of two normal distributions. When with such a test a small P results, it points to nothing except that there is a difference in means; if there is a substantial difference in means, such a P is likely to result.[2] Tests which Neyman and Egon Pearson call "universally most powerful" tend to be specific in this sense. When with such specific tests one has sufficient numbers, they become sensitive tests; in the terminology of Neyman and Pearson they become "powerful." I have suggested the generalization of which I feel only fairly sure, that highly specific tests are always equivalently statable in terms of an estimate and its confidence limits, and that it is better to state them in that way. There are being widely used in statistics today tests that are not good tests, in the sense mentioned, and these are the ones that I have been questioning. Examples are (1) the frequency chi-square test and (2) many applications of analysis of variance.

If the view is adhered to that what the small P of a test of significance does is to "disprove the null hypothesis tested" then not only will these latter tests be poor tests in being nonspecific or insensitive or both; they may be positively misleading. The case I cited of Professor Fisher's application of the technic of analysis of variance as a test for linearity is a case in point. This example was not brought out in isolation to point up anyone's individual blunder, but was part of a continued discussion of the pitfalls into which the pursuit of an erroneous conception of tests of significance could lead one. It was only one of many examples that I have published, mostly from my own experience. Professor Fisher applied[3] what he titled a "test of straightness of regression line" to some data of Hersh,[4] of eye facet number of Drosophila in relation to temperature, obtained a small *P* and concluded that the regression is nonlinear. He insists in his recent note that this is so, because this, he says, is the objective way of determining that it is not linear. Now as a naive desk neighbor of mine said when I read him the results, "If it isn't linear, what is it?" A plot of the points and a least square fit did not indicate any curvature that I or anyone else to whom I exhibited the graph was impressed with. If I were to predict facet number from a regression on temperature for these data, I should use a linear regression. On these grounds I wrote: "On inspection it appears as straight a line as one can expect to find in biological material." Professor Fisher thinks this judgment of the lack of curvature, instead of the reverse judgment based on the small *P* of his test, is obtuse and authoritarian. But I presented the evidence to the reader and I gave another explanation of the small *P*.

---

[2] J. Neyman and B. Tokarska, "Errors of the Second Kind in Testing 'Students,' Hypothesis, this JOURNAL 31: 318–326 (June) 1936.

I do not adhere to the doctrine that a small *P* disproves the "null hypothesis tested," in this case that the regression is linear. I have attempted to demonstrate in the article mentioned that, while this principle appears plausible, it is actually illogical and at variance with the usual mode of scientific demonstration. I do not think that a logician would support the proposition that "If an event known to occur rarely in hypothesis *H*, occurs, hypothesis *H* is disproved," and no validating experiments have been brought forward to support it either. In practical fields of investigation in which a test, *T*, is utilized, as for instance when a serologic reaction is used for the diagnosis of syphilis, validation of the test involves the determination of the frequency of *T* in (a) normal persons (null hypothesis) and (b) syphilitic patients (alternative). No one in this field would support such a test by demonstrating merely that if the population is normal, T occurs only rarely. Such an argument would appear absurd. What is wanted also, and indeed what is wanted mainly, is information as to how *often* T occurs if the population is syphilitic. It is this that determines the usefulness of the test for its purpose, the diagnosis of syphilis.

---

[3] R. A. Fisher, *Statistical Methods for Research Workers.* Ed. 7, Edinburgh, Oliver and Boyd, 1938, 356 pp.

[4] A. H. Hersh, "The Effect of Temperature upon the Heterozygotes in the Bar Series of Drosophila," *Journal of Experimental Zoology*, 39: 55–71 (February) 1924.

Why should the validity of statistical tests be immune to such critical appraisal as is required as a matter of course in any other field of application of tests? Analogous validation for statistical tests of significance would involve determining, first in what *alternatives*, and then how *often* the critical value of the test function occurs. Before the work of Neyman and Egon Pearson, no such inquiry was made for *any* statistical test of significance. Even now it has not yet been made for the tests that I have been questioning, despite raucous and repeated challenges from me to the mathematicians; for instance, it has not been made for all phases of the test of linearity of Fisher which we are discussing.[5] Examples of questions which would be answered by such an inquiry are: What is the probability of obtaining a value of $P \leqq 0.05$ with this test if (a) the abscissal variate x is not subject to variation, the ordinate variate y normally distributed with uniform variability σ_y, and the means of $\bar{y}_x$ follow some given nonlinear regression, (b) the regression is linear, the abscissal variate not subject to variation but the variability of y, σ_y is a given function, not constant, (c) the regression of means is linear, σ_y constant, but x not constant but subject to normal variation with σ_x?

I say the small P is evidence for any alternative hypothesis in which it will occur frequently (relative to the level of significance). In the instance in hand I suggested three[6] specifically and others generally: (1) curvilinearity as against linearity, (2) heteroscedasticity as against homoscedasticity, (3) variable temperature as against constant temperature, and (4) other conditions which will produce a small P. I gave as my own preference, in view of the fact that (to me) there was no confirmation of the nonlinearity, that the temperature was fluctuating. Examination of the figure suggests this explanation as especially likely if the temperature was variable at what was recorded as 23 or 25°C. or both. Professor Fisher says that, if I had read Hersh's article from which the data were obtained, I should not have committed the blunder which I made. I had read the article and it amply confirms my suggestions of possible alternative causes of the small P. There is definite evidence of heteroscedasticity in table 6 of the article, and as for fluctuation of temperature, here are Hersh's words (my italics):

---

[5] P. C. Tang, "The Power Function of the Analysis of Variance Tests," *Statistical Research Memoirs*. Edited by J. Neyman and E. S. Pearson, Department of Statistics, University College, London, December, 1938.

[6] At several points Professor Fisher achieves miscomprehension only with great effort. He says: "What Dr. Berkson fails to realise is that the judgment, from inspection, that the line appears as straight as one ought to expect, would have given no aid whatever towards discovering the cause of any real anomaly, whatever the cause might be, because that judgment in effect denies the evidence that any real anomaly exists." The fact is that it was I, not Professor Fisher, who suggested the multiplicity of possible meanings of the small P in this application. I said: " . . . a small P is evidence in favor of curvilinearity . . . , also . . . in favor of linearity plus heteroscedasticity . . . the temperature . . . subject to fluctuation . . . and . . . other conditions which . . . produce a small P relatively frequently . . . Which of these shall be taken to have been demonstrated . . . will have to be determined by other evidence, possibly other statistical tests." It is hard to imagine how I could have been plainer on this point. The possibility of a "remarkable discontinuity . . . in the neighborhood of 24°C.," not mentioned in Professor Fisher's textbook illustration, which is what I was commenting on, is provided for in the last part of my statement. We were discussing, however, if it requires recollection, not a test for *anything at all*, not just to see whether *something* was the matter. We were talking about a test for "straightness of regression line," that is a device by which we were to discern whether the regression was linear or curvilinear, and a specific application of it in which the regression had been found definitely to be nonlinear.

> Errors in this study may arise from fluctuating temperatures. The fluctuations were not the same in all cases; 15°, 17°, 25°, and 27° showed a variation of ±0.5°. 23° varied as much as 3°. The record of 29° showed a variation from 28.0° to 29.5°. Temperature readings taken of the food by a thermometer graded in tenths of degrees, showed a total variation of 1.4° in the 31° aquarium, and of 1.3° in the 21° aquarium. The 19° aquarium was more unsatisfactory and *varied sometimes as much as 5°*. In every case, however, an attempt was made to have the proper temperature during the facet reaction period, though *this may not have been accomplished* because the period may not be the same for the heterozygotes as for the homozygotes for which it has been determined. Nominal temperatures in accompanying tables and figures are to be considered with these remarks in mind.

EDWARD LEWIS DODD, 1875–1943

THE SUDDEN DEATH of Edward Lewis Dodd on January 9, 1943, at his home in Austin, Texas, following a heart attack, brought to a close a distinguished career. He had long been a member of the American Statistical Association and had just been elected a Fellow of the Association, though the news of his election did not reach Austin until shortly after his death.

Dr. Dodd was born in Cleveland, Ohio, November 1, 1875, the son of Heber and Emma Melvina (Drake) Dodd. He was educated in the Cleveland public schools; in Western Reserve University, from which he received the degrees of B.A. in 1897 and M.A. in 1901; and in Yale University, from which he received the degrees of M.A. in 1902 and Ph.D. in 1904. After serving as instructor in mathematics for two years at the State University of Iowa and one year at the University of Illinois, he came to The University of Texas as instructor in pure mathematics in 1907, and rose through the successive ranks of adjunct professor (1914), associate professor (1917), and professor (1923) of actuarial mathematics. In the year 1922–23, he taught at Williams College as visiting assistant professor of mathematics. On June 15, 1910, Dr. Dodd was married to Miss Alice Laidman of Austin, Texas. His wife survives him.

Dr. Dodd was a member of the honorary societies of Phi Beta Kappa and Sigma Xi and of various professional societies in addition to the American Statistical Association, including the American Mathematical Society, the Mathematical Association of America, the Econometric Society, and the Institute of Mathematical Statistics (of which he was a Fellow). He was a councilor of the Mathematical Association of America in 1920, and Associate Editor of the *American Mathematical Monthly* from 1918 to 1923.

His main fields of work were probability, statistics, and actuarial mathematics. During its early days he was a member of the Actuarial Advisory Committee of the Teachers Insurance and Annuity Association of America. He also gave much assistance in connection with the establishment of the Teachers' Retirement System of Texas.

Of the many honors which came to Dr. Dodd in recognition of his attainments in his chosen field, two will be mentioned here. He was chosen as the official representative of the Mathematical Association of America at the Seventh American Scientific Congress held at Mexico City, September 8–17, 1935. He was one of the two Americans who were invited to participate in the International Congress on the Theory of

Probability held at Geneva, Switzerland, October 11–15, 1937, and was the only American who actually did take part in it.

Dr. Dodd was an inspiring teacher, a distinguished scholar, and a master of productive research. A list of his publications, in addition to numerous reviews, abstracts, and minor papers, includes some forty items. But to those who knew him best, he will also be remembered for his sincere friendliness, his unfailing courtesy, and genuine kindliness, and his being always the Christian gentleman.

C. D. SIMMONS

The University of Texas

THE STUDY OF CYCLES

THE FOUNDATION for the Study of Cycles—a non-profit organization created to foster, promote, and conduct scientific research in respect to rhythmic and periodic fluctuations in any branch of science—announces the offering of a medal to the person who, during 1943, publishes the book or paper in this field that in the opinion of the judges is the most outstanding.

The judges are Dr. C. G. Abbot, Solar Physicist and Secretary of the Smithsonian Institution; Dr. Harold E. Anthony, Curator of Mammals and Dean of Scientific Staff, American Museum of Natural History; Professor W. C. Mitchell, Professor of Economics, Columbia University, and Director of the National Bureau of Economic Research; Professor V. C. Wynne-Edwards, Professor of Zoology, McGill University; and Professor Ellsworth Huntington, Professor of Geography and Climatology, Yale University (Chairman).

It may be in point to recite some of the reasons for assigning high importance to research in cyclic phenomena. First, in dealing with rhythm and periodicity the scientist is at the very heart of predictability. The power to predict accurately is the acid test of the degree of precision reached in any science. Medical science, for example, would become vastly more effective if we understood the degree to which various diseases are rhythmic in occurrence. Similarly, accurate long range weather predictions, which would be of inestimable value in agriculture and aviation, depend on a knowledge of the nature and causes of the more or less rhythmic variations observed now and in the past.

Second, the techniques for dealing with rhythms and periodicities are much the same in one branch of science as in another, but unless the subject is considered a field by itself, advance in methods in one discipline is often unknown in another until after the lapse of many years.

Third, the subject of rhythmic and periodic fluctuations is important because identical rhythms in unrelated phenomena and unrelated fields of science suggest possible interrelationships that might otherwise escape notice. For instance, long before the cause of tides was understood, the fact that the period of the tides is exactly half the period of the moon suggested that the moon must have something to do with the ebb and flow of the waters.

These are some of the reasons why a concerted study of cycles as a whole is important. There are of course other reasons in respect to individual branches of science.

We must all be careful not to confuse contributions to our knowledge of *interrelationships* with contributions to the knowledge of rhythms and periodicities *per se*. There may be an excellent correlation between two phenomena, and yet no rhythms may be involved in either. The direct concern of The Foundation for the Study of Cycles is only with rhythms and periodicities as such. We must also keep in mind that *denial* of the existence of a rhythm may be as important as affirmation.

In addition to awarding a medal, the Foundation will make awards to outstanding work in each branch of science. In these awards the judges will be assisted by the recommendation of an advisor appointed by a scientific society in the branch involved. For the American Statistical Association Dr. W. Edwards Deming, of the United States Bureau of the Census, has been designated.

The person who receives the medal and the persons who receive the awards will also be elected Fellows of the Foundation for the Study of Cycles.

Communications should be addressed to Professor Ellsworth Huntington, Chairman of the Committee on Awards, Hendrie Hall, Yale University, New Haven, Connecticut.

A CORRECTION

In the memorial note of Norman John Silberling in the March 1943 issue of this Journal, page 112, an error of omission was made. Dr. Silberling at the time of his death was Professor of Business Research at Stanford University.

BOOK REVIEWS

GLENN E. MCLAUGHLIN
Review Editor

Group Differences in Urban Fertility, by Clyde V. Kiser. Baltimore: The Williams and Wilkins Company. 1942. xii, 284 pp. $2.50.

This study, by a member of the Technical Staff of the Milbank Memorial Fund, forms an appropriate conclusion to an important series of investigations under private auspices directed toward the discovery and description of group differentials in fertility in the United States. Such investigations, on large scale and at high scientific level, were initiated by the same Fund a decade and a half ago. This study is the last of its kind, not because the problem can henceforth be neglected, but because its importance is now sufficiently demonstrated to be recognized as a matter of national interest and included in the agenda of the Federal Census. Private research in this field will henceforth be directed to related problems which are beyond the scope of census procedures, such as study of conditioning factors and motivation, and the elaboration of implications for social theory and policy.

It is apparent from Kiser's study and other investigations that the character of group differentials in fertility is undergoing an important modification. The general pattern of these differentials is not radically different today, but the monotonous iteration in the findings of various investigations is beginning to disappear. This is due, in part, to the development of more refined techniques; but it is also due in part to actual change in the phenomena themselves. As Ogburn has pointed out, differential fertility is in part due to lags in the diffusion of the family limitation pattern. So long as mere cultural distance from experimental science dominated the whole situation, the inverse relation of social status to fertility was fairly ubiquitous. In some cities, e.g., Boston, social gradients in fertility show little modification. This is true in general for all the large industrial cities in the Northeast. But in many parts of the country, group differentials in fertility show less uniformity and greater sensitivity to a variety of changing social conditions. Their accurate description will, therefore, require analysis of comparable data in different areas, under different conditions, and at different times.

Persons not concerned with the technical aspects of demography will generally await the publication of the forthcoming Census report on this subject for a description of the present status of group differences in fertility in the United States. This study by Kiser, however, explores some relations that will not be covered by the Census report. It also illuminates, in the course of the investigation, many other interrelations between occupational status, education, income, race, nativity, and locality.

Demographers will study this work primarily for its contributions to methodology. Many possible sources of bias in the interpretation of data on fertility and means of controlling such interpretations are fully developed by the author and need no further elaboration. It seems to this reviewer that the limitations of marital fertility rates based on statistics of births in any limited period and the bias inherent in the standardization of marital fertility rates if interpreted as measuring total marital fertility are not sufficiently emphasized by the author. A standardized marital fertility rate necessarily eliminates the influence of differences in the span of the exposure period, which is one of the major factors influencing group differentials in fertility. Also, as the author points out, age-specific marital fertility rates, unless controlled for order of birth, are appreciably influenced by the length of the preceding interval since marriage, and the mean length of this interval at any given age varies widely among different groups. Similar bias, in lesser degree, affects standardized frequencies of number of children ever born to married women. Age-specific fertility rates supply valuable information on factors affecting fertility. However, the most accurate index of total marital fertility is obtained from data on children ever born to ever married women in age classes near the end of the childbearing period, and preferably with classification by criteria such as race, place of birth, education, etc., which remain constant or fairly constant through the whole childbearing period. This suggestion is emphasized here because of its possible bearing on census procedures; but it is likely to encounter some resistence from tidy statisticians who dislike the rejection of any available data that may seem on the surface to be relevant to the problem at hand.

The predilection of statisticians for the use of all collected data and the impossibility of carrying through all possible breakdowns with a limited body of material may be responsible for the use of the ambiguous category, "colored." The behavior, as regards fertility, of the Negro population is radically different from that of Mexicans, Chinese, and Japanese. There were many of the latter groups in the western cities included in the National Health Survey on which this study is based. Limitation of the data for nonwhite populations to those for Negroes, except where separate treatment of other ethnic-cultural groups was possible, would have obviated some of the author's constant difficulties with this material, somewhat analogous to those of a scientist dealing with zoo data on the feeding habits of seals, snakes, and "quadrupeds." This suggestion also has possible implications for census procedure.

The treatment of pregnancy wastage, which is in general a very useful contribution, is handicapped, as the author points out, by the gross underreporting of abortions and its possible influence on apparent group differentials. For example, the relatively high frequency of pregnancy wastage in the case of women with college education in the professional group (found to be statistically significant) might possibly be due to fewer inhibitions in this group against telling the truth on this subject.

An adequate description of differences in fertility in the United States requires a vast body of data specifically designed for this purpose. This is clearly shown by the difficulties which a skillful scientist encounters in the attempt to use for this purpose data incidental to another study. Dr. Kiser and his predecessors and associates at the Milbank Memorial Fund are to be congratulated for blazing the trail, developing the necessary techniques, and demonstrating the possibility and significance of such studies. The Fund has already embarked on the still more difficult but no less important task of studying the conditions which determine such differentials and the possibilities of social control.

FRANK LORIMER

American University

Contraception and Fertility in the Southern Appalachians, by Gilbert Wheeler Beebe. Baltimore: The Williams and Wilkins Company. 1942. xii, 274 pp. $2.50.

When the millennium arrives with populations of optimum size maintained in every resource area of the world, the historians of that era should give due credit to the basic research in techniques and measurement of the control of human fertility which has been sponsored by the National Committee on Maternal Health and the Milbank Memorial Fund. Beebe's book follows closely the type of contribution made by Stix and Notestein in *Controlled Fertility*, with some further development in techniques for measurement and analysis of phenomena relating to contraception and fertility, and with adequate reporting of further experimental work.

Logan County, West Virginia, was the locale chosen for the three-year experiment reported. Within the county a group of women in the rural-nonfarm population, mostly wives of coal miners, were "enlisted" as contraceptive patients by a nurse working under supervision of local doctors. The technique advised was that of vaginal jelly alone with jelly and applicator supplied at no cost to the patient. For approximately 1,300 women who participated in the experiment, the chance of conception (as measured by number of second and subsequent conceptions per 100 woman-years of exposure to the risk of conception) was reduced by nearly one-half, from 73 to 40 for whites and from 82 to 45 for Negroes.

Such results form only the starting point of the detailed analysis presented. First is given an over-all review of the economic and social conditions of the Southern Appalachians of the sort that has been given before by regional sociologists and agricultural economists. Next follows a description of the project with particular attention to the problems of sampling involved. Like many other "experiments," this one began with scant concern for the representativeness of the individuals included in the study, but like few others, before the study was concluded, every effort was made later to explore their representativeness of rural-nonfarm women in the county. In fact, a major contribution to this line of research is Beebe's more careful consideration than has hitherto been evidenced of the importance of sampling procedures in the design of a study and of sampling errors in the drawing of conclusions from results.

The description of the group of women who participated in the experiment is interesting because they manifested in an exaggerated degree the traits common to the Southern areas of high fertility. Their mean age at first marriage was about seventeen and a half years, their admitted rate of premarital conceptions was 15 per cent for whites and 37 per cent for Negroes, their rate of first conceptions was such that 80 per cent had conceived within the first 10 months of marriage. From an auxiliary study which was made near the end of the experiment, it was found that about 32 per cent of the rural-nonfarm women of reproductive age in the county would accept contraceptive advice, but that these women because of higher fertility were responsible for about 50 per cent of the births of the county.

Hence the experiment demonstrates that similar programs on similar populations could reduce by about 40 per cent the fertility of a group of women who account for about 50 per cent of the county's births, or could reduce the total birth rate 20 to 25 per cent.

The conclusions and recommendations for the establishment of effective programs of contraceptive services again follow the Stix and Notestein lead of emphasizing the advisability of providing more than one technique to patients. This emphasis on the acceptability of a particular method to the patient rather than on the efficacy of a method under laboratory conditions has important implications for the planning of an actual clinic program.

MARGARET JARMAN HAGOOD

University of North Carolina

Public Health Statistics, by Marguerite F. Hall. New York: Paul B. Hoeber, Inc. 1942. xxii, 408 pp. $5.50.

In recent years there has been a virtual wave of new books on business statistics, social statistics, and general statistical methodology. Although the late Professor Pearl's third edition of *Medical Biometry and Statistics* appeared in 1940, the fact remains that new books devoted to vital statistics and biometry have been rare. This uncrowded condition of the field itself will probably ensure wide attention to Dr. Hall's *Public Health Statistics*.

The author states that this book "is intended for use as a text for a basic course in Public Health Statistics and as reference material in health departments." The introductory chapter contains a conventional but good discussion of history, principles, and limitations of statistical methodology. This is followed by three useful chapters on principles of computation and handling of mass data and tabular and graphical presentation of the materials. From this point the book can be said to fall into two major divisions, vital statistics and general statistics, although not so labeled. The vital statistics section includes Chapters V–XIV (pp. 88–259) respectively entitled population data; estimates of populations; standard millions; vital events; rates; mortality rates; vital statistics other than mortality; variation and trends in rates; ratios and index numbers; and construction of a life table. Chapters

XV–XXI (pp. 260–372), devoted mainly to general statistical methodology, are respectively labeled mass data—preliminary definitions; measures of central tendency; measures of variability; measures of relationship and trend; theory of probability, point binomial, normal curve; measures of reliability, and prerequisites to the application of statistical methods.

The author probably had good reasons for the tandem arrangement of the materials for vital statistics and those for general statistics, but the rewiewer fails to see what they were. If separation was necessary it would seem that the chapters on vital statistics should follow rather than precede those on general statistics. In the reviewer's opinion, however, better integration of the materials would be in order for a book on public health statistics. As one might expect, the present arrangement resulted in virtual duplication of the discussion in several instances.

The author avowed a conscious effort to prepare a book that could be used by people with little training in mathematics. With commendable simplicity of language and detail she leads the reader through the step-by-step procedures in specific computations. Dr. Halbert L. Dunn correctly states in the foreword, "She has discovered the knack of making probabilities, chi-square rates and index numbers into everyday county, city, and state health language." This characteristic will doubtless enhance the value of the volume as a handbook for use in health departments.

In the reviewer's opinion, however, the volume does not measure up so well as a textbook. For this purpose one would look for adequate discussion of underlying theory as well as for clarity in the exposition of technique. Unfortunately, the discussion of theory tends to be sketchy. It is possible that in writing the book the needs of the health department rather than those of the classroom were uppermost in the author's mind. Whatever the explanation, the author tended to devote much more attention to the mechanics of computation than to the meaning of the statistical measures considered. On several occasions the reviewer was somehow reminded of the lines, "Theirs not to reason why, Theirs but to do and die."

Despite the above characteristics there is ample evidence of the author's own well-balanced point of view regarding the place of statistical technique in public health work. She emphasizes strongly the limitations of statistical methodology and the necessity of knowing at first hand the characteristics of the basic data and the conditions under which they were collected.

CLYDE V. KISER

Milbank Memorial Fund

Refugee Settlement in the Dominican Republic, A Survey Conducted under the Auspices of the Brookings Institution. Washington: The Brookings Institution. 1942. xvi, 410 pp. $4.00.

In 1938 at an intergovernmental conference on the refugee problem, the Dominican Republic, alone among thirty-two nations, declared its willingness to welcome large numbers of refugee settlers. The limit was set at 100,000. This offer, the one bright spot in a gloomy picture, speedily led to the formation of the Dorsa, the Dominican Republic Settlement Association, a social agency financed from private funds, and to the establishment on the north coast of the Dominican Republic of the first refugee colony of under 500 persons at Sosua, a 26,000-acre estate offered to the Association by Generalissimo, now again President Trujillo.

This survey is limited to the activities and prospects of the Dorsa; it does not extend to other refugees, among whom Spanish loyalists are particularly numerous. Moreover, the survey confines itself strictly to economic aspects of the problem, omitting social, political, and other non-economic phases. Nevertheless, this book is far more than the study of a single local settlement. It is the most comprehensive study of the economic life of the people of the Dominican Republic known to the reviewer and as such constitutes a worthy companion volume to the Brookings Institution survey *Porto Rico and Its Problems*, 1930. The success or failure of the Dorsa and its protegés is inextricably tied up with the economic life of the Republic.

This economic survey of the Dominican Republic makes up the bulk of the study. It is preceded by three chapters dealing with the refugee problem in general and is followed by four chapters devoted to the work and problems of the Dorsa in particular.

The principal questions which the survey had to consider were: (1) What is the relation of the Dominican project to the refugee problem as a whole? (2) Is the Dominican Republic a suitable place for refugee settlement? (3) How may refugee immigrants make a living there? (4) How many settlers can the Republic absorb? (5) What will be the effect of the settlement project on the Dominican community?

Of these five questions the third and fourth are of most immediate and practical importance. The answers given are none too encouraging. Refugee immigrants may aspire to the level of peasants, i.e., agriculturists devoting most of their energies to subsistence farming with cash crop farming or primitive industry definitely supplementary and secondary. The Republic can absorb according to the majority of the survey staff from 3,500 to 5,000 persons and, according to one member, up to 10,000 refugee settlers.

Those familiar with the desperate conditions prevailing throughout the Caribbean area, the answer to question three can hardly be surprising. But the low absorption capacity of the Dominican Republic does come as a shock to those who compare the Dominican rural population density of 68 per square mile with that of Puerto Rico, Haiti, and Jamaica, which in the aggregate is over four times as high. The low arability coefficient, the fantastic rate of increase of the population, and the land requirements of the refugee peasants, which far exceed those of the native rural proletariat, offer the explanation.

It will cost not much less than $4,000 to settle a refugee on the island and of that he is expected to pay back about $1,600. At that rate the settlement of 5,000 "units" would require an outside net contribution of roughly 12 million dollars.

Apart from the chapter on the history of the Republic which reads like a dime novel, the book is written in a sober, almost sombre tone. It is an invaluable study of a local problem with global implications. It is required reading for all those wrestling with the problems of the post-war world.

The map of the Island of Hispaniola facing the title page might have been enhanced in usefulness by the insertion of names prominently mentioned in the text. The absence of an index is not a serious flaw in a survey carefully outlined and skillfully articulated.

ERICH W. ZIMMERMANN

The University of Texas

Consumer Instalment Credit and Economic Fluctuations, by Gottfried Haberler. New York: National Bureau of Economic Research. 1942. xix, 239 pp. $2.50.

This volume is the ninth of the studies in consumer instalment financing prepared under the financial research program of the National Bureau of Economic Research. In it data gathered in the previous studies are used as the basis for a theoretical analysis of the relations between consumer instalment credit and cyclical fluctuations in general business.

Consumer instalment credit both in the form of commodity credit and cash loans shows a pronounced cyclical pattern when measured by any one of the series used for this purpose, volume of new credit granted, amount of credit outstanding and "net credit change." The causes of fluctuations in instalment credit are found in both long run growth factors as well as cyclical factors. Cyclical fluctuations in consumer credit are caused more by cyclical shifts in demand than by cyclical shifts in supply. Responsiveness of demand to changes in the terms of credit is more dependent upon changes in the down payment percentage and contract length than upon changes in the rate of interest charged.

The introduction and expansion of consumer instalment credit in the 1920's had a stimulating effect upon the economy in the short run. Whether this development is to be regarded as favorable or unfavorable for economic stability in the long run depends upon questions of business cycle theory on which expert opinion is still divided. Haberler's conclusions are as follows. After the introduction and development to maturity of instalment credit as an institution it is still subject to cyclical fluctuations but these fluctuations are then the result rather than the causes of cycles in general business activity. However, consumer credit does act as an amplifier by intensifying the movements of the general cycle whether up or down. Yet this behavior is in no way different from that of credit in general.

In comparison with other influences upon general business activity consumer instalment credit was not very important during the decade 1929–40, much less important, for example, than the financial operations of the Federal Government. Yet consumer credit has a greater significance than this comparison would suggest when it is considered in relation to the development of particular industries and their effects upon the economy as a whole. Consumer credit as an institution should be maintained, but efforts should be made either to mitigate its cyclical fluctuations or to bring about an anticyclical pattern.

The publication of this book is timely because of the new regulations of consumer credit of the Federal Reserve System in the present emergency and the possibility that such regulations may be carried over into the post-war period.

James F. Cusick

Dartmouth College

Farm-Mortgage Credit Facilities in the United States, by Donald C. Horton, Harald C. Larsen and Norman J. Wall. Washington, D. C.: United States Government Printing Office. Miscellaneous Publications No. 478. United States Department of Agriculture. 1942. vii, 262 pp. 55 cents.

A more appropriate name for this publication would have been Farm-Mortgage Credit Facilities *of* the United States, Why they should be increased and How they can be increased. Nine of the sixteen chapters deal exclusively with federal credit agencies and the remainder consider the more general problems of debts, interest rates, etc., with but little mention, let alone description or analysis, of other farm-mortgage credit facilities. Yet, the following statement expresses a point of view which permeates the entire discussion.

> But the public purposes to be achieved through extensive participation of the Government in farm-mortgage financing may be so essential as to make the disadvantages negligible. It is possible that the results expected of a coordinated Federal mortgage-credit service are such that the profit motive of private financing alone cannot be expected to accomplish them. For example, public policy operating through Federal mortgage-credit channels might be made to serve as a beneficial form of public action designed, along with other action programs, to assist agriculture in adapting to changing economic conditions. Considerations of this character may justify the sacrifice of certain of the advantages to farmers of private credit facilities. (Page 170.)

The above can hardly be called a conclusion because it is not based on any evidence in the text. It is not stated as an assumption on the basis of which further argument is advanced. It does, however, express a philosophy which is evidenced in numerous parts of this work. The reviewer has no wish to quarrel with the authors for this point of view. He does wish to point out, however, that evidence is not presented to support it other than evidence which indicates that the Federal Government did expand its credit facilities in the depths of depression. The fear expressed is that unless these agencies adapt themselves to present economic conditions they may not be available to serve should similar catastrophies arise in the future. Then, of course, the happy thought occurs that such a calamity may be avoided if "wise" credit policies are followed.

The authors maintain that "if the Federal agencies are to be major sources of farm-mortgage credit in periods of prosperity as well as in depression, it may be necessary to modify these institutions further so they can serve not only as refinancing and financial-adjustment agencies but also as continuous sources of both purchase-money and operating credit."

Though the authors may seem to the reader to be advocates of "socialized credit" they do advocate the sound policy of modifying by improvements existing Federal farm-mortgage credit facilities to meet the varied needs of agriculture under varied conditions. As a basis on which such modifications may be made there is presented a wealth of information, tabulated in a very orderly manner and unusually well illustrated.

Part I discusses the historical background of the farm-mortgage credit situation for the period 1910 to 1940. Though the general theme is much the same heard so many times, particularly during the last decade, its concise recording with suitable statistics is a useful adjunct to the text.

Part II is particularly good in its description of Federally sponsored farm-mortgage loan agencies. The organization and development of each agency is described and its source of funds, terms and conditions of loans explained.

The first two parts carry out very well the intentions of the authors, namely "to serve a principal purpose of providing a compact summary and analysis of factual material on farm-mortgage credit heretofore available only in scattered sources."

Part III is concerned with current problems and issues in farm-mortgage credit policy but as indicated in the introduction "more emphasis is placed upon identification of problems and issues, and upon exploration of their interrelationships in the light of alternative public objectives than upon formulation of specific recommendations."

The publication merits careful study by all borrowers and prospective borrowers and by lenders, both public and private, at this particular time when foundations are being built for the economic structure within which we must all live for the next decade or two.

JOHN B. SCHNEIDER

University of California, Berkeley

Business Behavior 1919–1922, by Wilson F. Payne. Chicago: The University of Chicago Press. 1942. xi, 215 pp. $1.50.

This book must be regarded as timely, since many people are now interested in accounts of past experiences of the postwar development of business which may be used for the purpose of forecasting the future development of business after the present war. The approach of the author to his subject is more than empirical; it may be qualified as "atheoretical." In the introductory chapter, he complains of too many contradictory theories of business fluctuations and points to the danger of selecting some one hypothesis as a basis for analysis of data, since this selection of itself determines what kind of data is considered as important in such an analysis.

The author follows an industry approach to his subject. Several industries were selected for survey, most of which are divided into a number of integral segments. He analyzes leather and shoes, textiles, iron and steel and their principal users, automobiles, and building and allied industries. The last group is not included for close examination, but simply to complete the general picture of the period. These surveys are preceded by a short analysis of the financial situation during the period under study. The author recognizes that this aspect of business development during 1919–22 has received keen examination by other observers, but he found it advisable to review it in order to provide necessary points of reference.

The author thoroughly worked numerous statistical data, which are summarized graphically in eleven charts, each including some 15 to 20 curves. These charts represent the behavior of the various indexes characteristic of the industries surveyed in the book. But they are not interpreted in the text. This may be regarded as a shortcoming from the point of view of the general reader, since comparisons and interpretations of these charts are not easy problems. Only a specialist may do so advantageously. In the opinion of the reviewer, this omission, which presumably was done for economy of space, may be partly responsible for the fact that Payne's book does not leave the reader with an impression of the behavior of business as a whole. Although the interactions between segments of individual industries are sufficiently analyzed in separate chapters, little attention is given to a systematic analysis of the interactions between the major branches of industry during the 1919–22 cycle.

The last chapter gives a recapitulation of developments, but it occupies only nine pages, and its conclusions do not utilize fully the rich material supplied in the separate chapters. The lack of sufficient analysis of interconnections may result from the atheoretical approach of the author. He purposely avoids using any of the analytical tools elaborated by the theory of the business cycle, as if he is afraid of falling in the tenets of some one theory. But his statement that the list of industries selected for analysis "is confined to a representation of those of which short industrial depressions are peculiarly composed" presupposes some theoretical conception of the business cycle, which he does not formulate explicitly. By excluding the food industries and agriculture, he disposes of at least one group of the theories of the business cycle. The exclusion of agriculture is not consistent, however, since in his analysis of the textile industries he was obliged to pay much attention to variations of cotton crops and his analysis shows how important the quantitative and qualitative characteristics of these crops during the years under study were for the explanation of the business situation in the cotton industry.

The extreme empiricism of the author, however, may be justified by the unsatisfactory state of the theory of the business cycle. In its present situation, the theory may be regarded more of a barrier than a help to a theoretico-empirical approach of the subject, and may contribute to extreme empiricism of even those investigators who are conscious of the dangers of extreme empiricism, as is the author.

The book may be used advantageously in advanced courses on the business cycle, particularly in special seminars, though the general reader will hardly get a general picture of the business behavior during 1919–22. Those interested in the behaviors of separate industries during that period will find the book interesting and useful reading.

V. P. TIMOSHENKO

Food Research Institute
Stanford University

Commodity Statistics, The 1942 Commodity Year Book. New York: Commodity Research Bureau. 1942. 294 pp. $5.00, and *Commodity Year Book.* New York: Commodity Year Book. New York: Commodity Research Bureau, Inc. 1942. 413 pp. $7.50.

The volume on *Commodity Statistics* is the fourth of an annual statistical series. Descriptive materials on the background, processing, and marketing of commodities have been omitted from the 1942 volume. Such information covering all the commodities listed in the statistical volume and in addition virtually all of the important crude and semi-finished products used in the national economy are presented in the *Master Edition, Commodity Year Book.* The latter, therefore, is a commodity encyclopaedia designed for permanent reference use.

Commodity Statistics presents data in 61 individual or group sections for the 75 important crude and semi-finished materials covered in previous statistical editions. Statistics reported are for commodities generally classified as animal and animal products, beverages, forest products, fuels and electric power, metals, nonmetallic minerals, vegetable foodstuffs, vegetable oils and fibers, and such specialties as rubber, silk, and tobacco.

Data presented in each section vary in length from 1 or 2 pages to 16 pages depending on available information and the importance of the commodities. The sections generally comprise statistics on consumption or sales, production, prices, exports and imports, and stocks. Frequently statistics are given on such subjects as yields, taxes, kinds of uses, and sizes. All the statistical tables are simply arranged and easy to read. Footnotes are given but in some tables, particularly those involving estimates and comparability of a series of data, technicians may desire more detailed explanations.

The data have been compiled from a variety of recognized sources such as Government agencies, Boards of Trade, commodity exchange bureaus, and other institutions. Source references to the pertinent agencies and institutions are printed at the foot of each table. This source information could be more helpful to statistical workers if it also included references to the names of the current publications of the particular agencies. Technicians using the volume may find it necessary to look up available current data.

Information in the 1942 Year Book has been somewhat clouded by war conditions. Only data that have been officially released by governmental and quasi-official sources as not being detrimental to the war effort have been published. Consequently, statistics for strategic and critical commodities generally are not shown after the first 9 to 11 months of 1941.

The *Master Edition* is composed of separate articles on over 800 different crude and semifinished commodities. These articles contain brief descriptions of available information on such subjects as physical characteristics, methods of production, marketing and transportation methods, comparative prices over recent years, principal types and grades, possible substitutes, and government regulations. An attempt was made to use the same general outline for all commodity articles, but variations are frequent because such data for many commodities are lacking in part or otherwise do not warrant such specific treatment. Amount of material presented for each commodity varies from part of one page for the less used commodities up to 16 pages for a group of commodities such as plastics. Notable features of this volume are its brief historical and economic reports on rubber, plastics, and many common and rare chemical and vegetable commodities.

The two volumes should be valuable to statisticians, investigators, and research workers, because extensive statistical and informational data for so many important commodities are to be found in one set of reference books. It is hoped that the editors will find it possible and necessary to expand the coverage of their statistical volume to include some data for all commodities described in the *Master Edition*.

Oscar L. Endler

National Resources Planning Board

Make or Buy, by James W. Culliton. Boston: Harvard University, Bureau of Business Research. 1942. iv, 130 pp. $1.50.

Every business purchase, the author points out, involves at least a tacit decision that it is more desirable to buy the article in question from an outside supplier than to produce it within the organization. Conversely, every production schedule indicates at least a tacit decision that each item involved can more satisfactorily be made by the company than purchased outside.

While there are limits to the field of such choices, the author believes that too often a decision to make or buy a particular item, regardless of its original merit, continues to stand as an established policy long after changing conditions call for further review. Obviously, few companies will find it economical to make their own lead pencils or stationery; a specialty manufacturer will not seek to put himself out of business by buying the finished articles which provide his primary manufacturing revenue. Between such limits, however, lies a wide range of optional items which could be either made or bought.

Cost will be the principal factor determining such a decision, the author emphasizes. He advocates no one method of determining costs, pointing out that the facts will vary from case to case in the matter of overhead involved in producing a particular part, in the length of tool life, and in optional uses of existing facilities. Savings or increases in purchasing expense and storage and handling costs may or may not be capable of determination and allocation for individual items. In the analysis of a number of case studies, the author points out that all costs that can be clearly allocated should be, and that the decisions reached in some cases might have been reversed had a different analysis of cost been made. The final decision, even where cost is concerned, depends to a large extent upon the exercise of business judgment.

Further factors which will influence "make or buy" decisions, particularly where the cost element is not decisive, are control over quality, control over scheduling of production, and risks beyond the company's control. The author questions whether these factors deserve as much weight in decisions not to buy as they often receive, and cites cases to support his position.

Readers in search of clear-cut generalizations as to the preferability of making or buying will be disappointed. The analysis and the wealth of case material from a variety of manufacturing operations will, however, make the book of interest to individuals reviewing their own company policies. With many companies facing postwar reconversion and a fresh start in their regular peacetime operations, the book is particularly timely.

CHARLES E. YOUNG

Westinghouse Electric and Manufacturing Company

Wool and the War, by Hiram S. Davis. Philadelphia: University of Pennsylvania, Wharton School of Finance and Commerce, Research Report No. 6. 1942. xi, 107 pp. 50 cents.

This publication presents a clear picture of our war supplies problems for wool textiles. The author carefully weighs the wool supply prospects compared with probable needs for the war period. The report analyzes stock on hand and prospective production and importation of raw wool, prospective production of substitute materials, and the country's capacity for producing wool textiles as measured by mill facilities and trained labor. Wool needs are appraised in terms of prospective military and civilian requirements through 1944.

The author concludes that, although our present and prospective supply of raw wool may not be enough to support civilian consumption at its peacetime rate without encroaching upon military reserves, there is no foreseeable shortage of substitute fibers ahead. Continuance of the use of raw wool for civilian products at its current restricted level (the conservation order of the WPB) has limited civilian use to about 44 per cent of annual peace-time consumption, and the difference between this volume and usual peace-time consumption the author contends can easily be made up by using in apparel production more of the wastes from wool processing, more wool fiber recovered from tailor's clips and old rags, more hair fibers, and finally more blends of wool and rayon or cotton, or of cotton-warp wool goods.

The author warns that this plentiful supply of substitutes should not mean that consumers can expect to buy wool textile products at the high level of 1942 without sooner or later creating serious shortages. This warning note seems very appropriate in view of the prospective increase in purchasing power in 1943 and 1944, coupled with the expanded rationing program of non-clothing items which concentrates more consumer buying pressure on clothing items.

The report concludes that the maximum military and civilian demand for wool textile products, which now seems in prospect, could be met from existing mill facilities if they were operated on a three-shift rather than on a two-shift basis. Even with two shifts of 40 hours per week, however, the report concludes that the industry could handle a military business equivalent to the total of 1942 and a civilian output of apparel, blankets, and woven felts very near the level of 1939.

According to the author both materials and productive facilities should be adequate for prospective military and civilian demand. It should be remembered, however, that this report was written during the first half of 1942 and events which have occurred since that time have changed significantly some of the prospects particularly on the demand side. For example, the author bases his estimates of American supply on a maximum of five million men in the armed forces at the end of 1942, seven million by the end of 1943, and nine million by the end of 1944. Most recent estimates indicate at least eleven and a half million will be in the armed forces by the beginning of 1944, or shortly thereafter. This increase in demand, coupled with continued heavy buying of wool textiles, makes the wool supply picture less bright. Partly offsetting this greater demand situation is the record 1942 wool production of 459,073,000 pounds, which is an increase of nearly sixty-nine million pounds over the 1941 output. However, the author concludes that "the sheep expansion program does not offer significant results either from the point of view of the immediate emergency or from that of nationalism." In support of this conclusion, the author cites the failure of higher wool prices during World War I to result in any appreciable increase in sheep numbers. Despite the tripling of prices, the domestic clip in 1919 was only five per cent above that of 1914, and the wool clip would have shown an actual decline had it not been for expansion in Texas production. It is significant that sheep numbers in the United States at the beginning of 1943 are about three per cent below those a year earlier.

The report is well buttressed with ample supporting factual information and a thorough knowledge of the wool textile industry. Moreover, the author is aware of developing events because he ends the report with a warning note that rationing wool textiles may be necessary not because wool textile supplies may fall short of ordinary needs, but because unspent purchasing power may push demand up to extraordinary levels. Also he points out appropriately that neither the discouragement nor rationing of consumer purchases really solves the problem of existing purchasing power, and that some steps must be taken to remove at least a portion of existing consumer purchasing power from the markets. Heavier taxation, enforced savings, and related fiscal controls are essential if purchasing power is to be kept down to a point where consumer sales will not create severe shortages of most wool textile materials. The implications of these requirements on national fiscal policy and the OPA rationing program should be obvious.

R. R. RENNE

Montana State College and
Office of Price Administration

Wartime Rationing and Consumption. New York: Columbia University Press. League of Nations Publication. 1942. II. A. 2. 87 pp. $1.00.

As its preface states, this short book "is concerned mainly with the impact of war controls and rationing on consumption and standards of living in countries for which information is available."

Its first brief chapter treats of methods and principles of consumption control and of direct consumer rationing. An important observation is that wartime governmental rationing systems " . . . contain the elements of a distributive system in which consumption is guided not so much by individual purchasing power as by human wants." It will be interesting to see whether postwar developments nurse or strangle this growing plant.

The long second chapter presents statistical evidence of food rations in many countries as of April 1942. For understandable reasons the sources of data cannot be given, and therefore the results cannot be checked by reviewers. The wartime history of changes in rations is touched briefly, and comparisons are drawn between the 1942 rations and prewar consumption. The value of this compilation and comparison, which represents immense labor, goes almost without saying; it constitutes the most comprehensive survey presently available. Its value is not vitiated by the following comments: Rumanian (Bucharest) rations of sugar, meat, and fats look incredibly high. The conclusion that Italy ranks with France, Belgium, Norway, Finland, and Spain as having rations so abnormally low as to imply prevalence of both undernourishment and malnutrition is difficult for the reviewer to reconcile with such advices as seep out of the country. The section on "Worker's Family Consumption" carries little conviction so far as concerns comparisons of rations with prewar calorie consumption as shown by budget surveys, because the range of calorie consumption indicated by the surveys was incredibly large. Acceptance of these surveys, and of various nutritional standards advanced by the League of Nations, seems somewhat uncritical.

The very brief third chapter gives a worldwide survey of wartime rationing of clothing, housing and household essentials, miscellaneous necessities and services, and non-essentials. The final chapter seeks to measure the effects of wartime spending restrictions upon the aggregate volume of consumption in seven countries, especially the United Kingdom, Germany, and the United States.

Readers interested in finding a guide for American policy may question the emphasis upon the desirability of equity of food distribution in a full-fledged war economy. They may also feel a lack of discussion of the labor costs of consumer rationing, of the impact of rationing on the distribution system, of the circumstances of general or only of particularized food shortages that justify food rationing, and of the difficulties of holding farmers within food-ration allotments.

The subject matter—principles and practices alike—lends itself to continuing inquiry and development. It is to be hoped that supplementary publications will appear.

M. K. BENNETT

Food Research Institute
Stanford University

Statistical Methods Applied to Agricultural Economics, by Frank A. Pearson and Kenneth R. Bennett. New York: John Wiley & Sons, Inc. 1942. vii, 443 pp. $4.00.

This is a valuable book for the practical agricultural economist. It combines in a single volume nearly everything of a practical nature which the average student of agricultural economics needs to know about statistical method. The authors have illustrated the various statistical techniques with numerous examples taken from practical problems largely in the fields of farm management, marketing, and prices. Statistical theory is subordinated to its practical application to real problems, but the student should easily get the theory from study of the specific examples to which it is applied. The language of the text is fairly non-mathematical. The interest and attention of the agricultural economist reader will be held by the wide variety of well chosen practical problems which illustrate the statistical procedures discussed.

The general arrangement of the book follows the usual order of statistics textbooks: measures of central tendency, variation, association, and reliability. It is divided into 22 chapters. Nearly every chapter has a very helpful section near its end entitled "Uses," and/or "Comparison," in which the authors summarize in a non-mathematical and practical manner the meat of the chapter.

Two excellently prepared chapters on the tabular analysis of relationships cover a field which has been neglected entirely too long in agricultural statistics books. This method has been and will continue to be the most widely used method of analyzing relationships. Perhaps its general usage and apparent simplicity have prompted statistical writers to dismiss it lightly. Yet it is well for the student of agricultural economics to know its uses and limitations. The manner in which the authors develop the treatment of tabular analysis provides an easily understood introduction to the theory of correlation in the following chapter.

Regression and correlation are treated in a more understandable manner than is true in many textbooks. There is considerable material of a practical nature in these chapters that is frequently omitted from books of this kind.

Some readers may feel that the authors have placed too much emphasis upon the analysis of data already collected, and too little emphasis upon the theory of sampling and the planning and collection of an adequate and representative sample. Agricultural economists too frequently follow some variation of the "hit or miss" method in the collection of sample research data, and then go to great pains to be "accurate" in its summarization. It is essential that agricultural economics students and researchers alike be made increasingly aware of the importance of proper sampling technique.

Other readers may notice the omission of an introductory chapter explaining the general relationship of statistics to science and research. Statistics are a very useful tool, and it may seem that any statistics textbook should attempt to orient the beginning student in an appreciation of the contribution that statistics can and should make to the advancement of knowledge.

E. L. BUTZ

Purdue University

Fundamental Statistics in Psychology and Education, by J. P. Guilford. New York: McGraw-Hill Book Company, Inc. 1942. xi, 333 pp. $3.25.

The intention of the author in preparing this book is to present the basic principles and methods of statistics to the beginner in the subject. The treatment is essentially non-mathematical in character, although mathematical symbolism is freely used.

No attention is given to methods of collection and arrangement of data. The major emphasis is upon methods of analysis and the interpretation of the results of analysis. Early chapters deal adequately with measures of frequency distributions, and the properties of the normal curve. The remainder of the book is devoted to the problems of sampling, measures of reliability of samples, and correlation procedures.

Perhaps the outstanding feature of the book is the uniform clarity of presentation. There are certain points at which the reviewer feels that the fetish of brevity may have been pursued at the expense of clarity and completeness, but such lapses are negligible in comparison with the overall precision of expression and adequacy of treatment.

This matter of brevity is most prominent in the last two chapters of the book. The chapter on multiple and partial correlation would be more satisfying if an explanation were included of the meaning of multiple and partial regression equations. Some explanation is also needed of why the successive steps are taken in the Doolittle Method of solution. The chapter on reliability and validity of tests is an excellent critical discussion of the theoretical problems involved in evaluating test results. It seems to imply more practical experience with tests than the beginning statistical student would ordinarily be presumed to possess.

At the end of each chapter a set of exercises is appended. These exercises are well designed to reinforce the student's understanding of the principles set forth in the chapters, and are adequate for homework assignment or for a laboratory of moderate length.

The attitude of students toward courses in statistics has in many cases become extremely adverse because the authors of textbooks have succeeded only in introducing mystery into the subject. This book takes the mystery out of the subject by its clear and rounded exposition, its judicious choice of illustrative examples, and its general adherence to the point of view of the beginning student. The book is indeed a valuable teaching medium.

MARTIN A. BRUMBAUGH

University of Buffalo

An Outline of Probability and Its Uses, by Maurice C. Holmes. Minneapolis: Burgess Publishing Company. 1936. viii, 119 pp. $1.50.

Four decades of experience in a great industrial organization justify this reviewer in stating that the *Outline* by Professor Holmes should be in the hands of every engineer and statistician unfamiliar with the theory of probability. Moreover, the *Outline* offers an excellent résumé for one already acquainted with the subject

Commendation is called for by the author's choice of material and the order of its presentation. The diagrams throughout the text will appeal particularly to the engineer.

In this time of stress, when our thoughts should be concentrated on winning the war, the above should suffice as a review. However, the following comments are added in anticipation of a second and other editions, and as proof that the reviewer has read the first edition from a to z.

1. On reading page 7, where mutually exclusive events are dealt with, the novice may interpret "the probability that *at least* one will happen" as meaning that "not less than one will happen." It might be clearer to say "some one of the events" or "some one or other of the events will happen."

2. On page 10, no doubt the author is historically correct in his statement that equation (9) is "called the binomial law from the fact that the term C_n^m is equal to the coefficient of the nth power of x in the binomial expansion of $(x+y)^m$." In our industry we apply the name "binomial law" to equation (9) because its entire right hand term is a term in the expansion of the binomial $(p+q)^m$, wherein $q=1-p$.

3. On page 99, under the caption "Symbols," the equation

$$a+b=\text{"}a \text{ or } b\text{"}$$

might better be written

$$a+b=\text{"}a \text{ or } b \text{ or both."}$$

One unfamiliar with symbolic logic may interpret "*a* or *b*" as meaning "*a only* or *b only*."

The above comments in no way reflect on the merits of the *Outline*. The exposition goes straight to the point without any lost motion. Moreover, it is entirely free from facetious phrases, a type of phrase which disfigures many otherwise excellent works.

E. C. MOLINA

Seven-Place Values of Trigonometric Functions, compiled by J. Peters. New York: D. Van Nostrand Company, Inc. 1942. $7.50.

Those acquainted with the well known tables *Siebenstellige Werte der trigonometrischen Funktionen* compiled by Dr. J. Peters of the Kgl. Astronomische Rechen-Institut, Berlin, and published in 1918, will welcome this photo-offset of those useful tables together with the English translation of the Introduction. The Alien Property Custodian has undoubtedly performed a valuable service in making these tables generally available at this time.

The need for increasing the speed of computations and the increased use of electric calculating machines in astronomical, geodetic, and optical calculations has emphasized the value of accurate natural trigonometric functions. These tables admirably fill that need. The use of decimal parts of degrees is logical in simplified calculations. Since these tables are photo-offsets they will not be subject to typographical errors as new type-set tables would be. This reviewer has used the original tables with the electric calculator for over a decade finding them more satisfactory than any others currently available.

These tables which give the sine and tangent (cosine and cotangent) in thousandths of a degree and the value of the function to seven places are based upon those of Gellibrand's *Trigonometria Britannica*, (Goudae 1633) which were computed for every hundredth of a degree of the quadrant. The functions in the present tables were interpolated for the thousandth of a degree, except for the tangents of angles from 89° to 90° where the values were calculated. The possible maximum error is given as ± 0.65 unit in the seventh decimal place of the sine function and as ± 0.85 unit in the seventh decimal of the tangent function. In general the errors are very much less than these and interpolations may be made for a ten-thousandth of a degree with considerable accuracy. Conversion tables of minutes and seconds to decimal parts of degrees and of degrees to time (and vice versa) are given at the end of the book. The tables are arranged for convenient use.

KENNETH N. OGLE

Dartmouth Eye Institute
Hanover, New Hampshire

PUBLICATIONS RECEIVED*

Barger, Harold, and Landsberg, Hans H. *American Agriculture, 1899–1939: A Study of Output, Employment and Productivity.* New York: National Bureau of Economic Research. 1942. xxii, 440 pp. $3.00.

Bennett, M. K., Farnsworth, Helen C., and Peirce, Rosamond H. *Wheat in the Third War Year: Major Developments, 1941–42.* Stanford University, California: Wheat Studies of the Food Research Institute. Vol. XIX, No. 3. 1942. 35 pp. $1.00.

Bernays, Edward L. *The Future of Private Enterprise in the Post-War World.* New York: Counsel on Public Relations. 1942. 11 pp.

Bowley, A. L. *Studies in National Income.* Cambridge: The University Press, New York: The Macmillan Company. 1943. ix, 255 pp. $3.50.

Bowman, Mary Jean, and Bach, George Leland. *Economic Analysis and Public Policy.* New York: Prentice-Hall, Inc. 1943. xix, 935 pp. $4.50.

Clark, Fred E. *Principles of Marketing.* New York: The Macmillan Company. 1942. xx, 828 pp. $4.50.

Culliton, James W. *Make or Buy.* Boston: Harvard University, Bureau of Business Research. 1942. iv, 130 pp. $1.50.

Davis, Joseph S. *New International Wheat Agreements.* Stanford University, California: Wheat Studies of the Food Research Institute. Vol. XIX, No. 2. 1942. 58 pp. $1.00.

Fabricant, Solomon. *Employment in Manufacturing, 1899–1939.* New York: National Bureau of Economic Research. 1942. xix, 362 pp. $3.00.

Fisher, R. A. *The Design of Experiments.* Edinburgh: Oliver and Boyd Ltd. 1942. Third Edition. xi, 236 pp. 12 shillings 6 pence.

Frickey, Edwin. *Economic Fluctuations in the United States.* Cambridge: Harvard University Press. 1942. xxi, 375 pp. $5.00.

Gras, N. S. B. *Harvard Co-operative Society, Past and Present.* Cambridge: Harvard University Press. 1942. xi, 191 pp. $1.50.

Hardy, Charles O. *Do We Want a Federal Sales Tax?* Washington: The Brookings Institution. 1943. 47 pp. 25 cents.

Hoyer, Louis P., and Hay, Charles K. *Services to the Orthopedically Handicapped.* Philadelphia: Walther Printing House. 1942. xv, 115 pp.

Hurlin, Ralph G. *Salaries and Qualifications of Child Welfare Workers in 1941.* New York: Russell Sage Foundation. 1943. 27 pp. 25 cents.

International Labour Office. *Year Book of Labour Statistics.* Montreal: International Labour Office. 1943. 222 pp. $3.00.

Kurtz, Russell H., Ed. *Social Work Year Book, 1943.* New York: Russell Sage Foundation. 1943. 764 pp. $3.25.

League of Nations. *Economic Fluctuations in the United States and the United Kingdom, 1918–1922.* New York: Columbia University Press. 93 pp. $1.50.

Lombardi, John. *Labor's Voice in the Cabinet.* New York: Columbia University Press. 1942. 370 pp. $4.00.

Magill, Roswell. *The Impact of Federal Taxes.* New York: Columbia University Press. 1943. ix, 218 pp. $3.00.

Mendershausen, Horst. *The Economics of War.* New York: Prentice-Hall, Inc. 1943. xiv, 390 pp. $3.35.

Merwin, Charles L. *Financing Small Corporations in Five Manufacturing Industries, 1926–36.* New York: National Bureau of Economic Research.

---

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. Some items of minor interest to statisticians have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications which are to be found in the *Revue de l'Institut Internationale de Statistique*, *Journal of the Royal Statistical Society*, *American Economic Review*, *Population Index*, *Transactions of the Actuarial Society of America*, *The Record of the American Institute of Actuaries*, and *Sankhyā, The Indian Journal of Statistics*.—Editor.

Studies in Business Financing. 1942. xvi, 172 pp. $1.50.

Metz, Harold W. *Is There Enough Manpower?* Washington: The Brookings Institution. 1942. 25 pp. 25 cents.

Mitchell, Wesley C. *Wartime 'Prosperity' and the Future.* New York: National Bureau of Economic Research. Our Economy in War, Occasional Paper 9. 1943. 40 pp. 35 cents.

Moulton, Harold G., and Schlotterbeck, Karl. *Collapse or Boom at the End of the War.* Washington: Brookings Institution. Pamphlet No. 47. 1942. 40 pp. 25 cents.

National Bureau of Economic Research. *Income Size Distributions in the United States, Part I.* New York: National Bureau of Economic Research. 1943. xxvi, 131 pp. $1.00.

Nourse, Edwin G. *Democracy as a Principle of Business.* Washington: The Brookings Institution. 1942. v, 24 pp. 25 cents.

Nourse, Edwin G. *The Timing of Price Changes.* Washington: The Brookings Institution. 1943. ix, 35 pp. 25 cents.

Perkins, John A. *The Role of the Governor in Michigan in the Enactment of Appropriations.* Ann Arbor: University of Michigan Press. Michigan Governmental Studies, No. 11. 1943. ix, 197 pp. 75 cents.

Robinson, Sophia M. *Jewish Population Studies.* New York: Conference on Jewish Relations. 1943. xvi, 189 pp.

Saulnier, Raymond J., and Jacoby, Neil H. *Accounts Receivable Financing.* New York: National Bureau of Economic Research. 1943. xv, 157 pp.

Schultz, Theodore W. *Redirecting Farm Policy.* New York: The Macmillan Company. 1943. vii, 75 pp.

Shepard, E. Fenton, and Wood, William B. *The Financing of Public Schools in Michigan.* Ann Arbor: University of Michigan Press. Michigan Governmental Studies, No. 13. 1942. 139 pp. 60 cents.

Shirras, G. Findlay, and Rostas, L. *The Burden of British Taxation.* New York: The Macmillan Company. 1943. xiii, 240 pp. $3.00.

Tamagna, Frank M. *Banking and Finance in China.* New York: Institute of Pacific Relations. 1942. xxi, 400 pp. $4.00.

Van Vleck, George W. *The Panic of 1857.* New York: Columbia University Press. 1943. ix, 126 pp. $1.50.

Wells, Ralph G., and Perkins, John S. *New England Community Statistical Abstracts.* Boston: Bureau of Business Research, Boston University College of Business Administration. 1942.

Weston, John Frederick. *The Economics of Competitive Bidding in the Sale of Securities.* Chicago: The University of Chicago Press. 1943. ix, 51 pp. $1.00.

Whittlesey, Charles R. *The Banking System and War Finance.* New York: National Bureau of Economic Research. Our Economy in War, Occasional Paper 8. 1943. 53 pp. 25 cents.

Zapoleon, Marquerite W. *Community Occupational Surveys.* Washington: United States Government Printing Office. Occupational Information and Guidance Series No. 10. 1942. v, 199 pp. 25 cents.

Government Publications

Department of Commerce. *Survey of Business Research Projects at Universities.* Washington: Bureau of Foreign and Domestic Commerce. 1942. v, 188 pp.

State of New York. *Careers in New York State Government.* New York: Department of Civil Service. 1942. 62 pp.

A Economia Catarinense. Santa Catarina: Departamento Estadual de Estatistica. 1942. 36 pp.

Anuario Estadistico de la Republica Dominicana del Ano 1941. Santiago: Editorial el Diario. 795 pp.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

SEPTEMBER, 1943 · $1.50 PER COPY · $6.00 PER ANNUM · VOL. 38 · No. 223

CONTENTS

ARTICLES

Observations on the Cost of Living Index of the Bureau of Labor Statistics . . . . LAZARE TEPER 271

Notes on Mr. Teper's Observations . . . . ARYNESS JOY WICKENS AND FAITH M. WILLIAMS 285

Analysis of Variance for Percentages Based on Unequal Numbers . . . . W. G. COCHRAN 287

"Retro" Charts . . . . . . KARL KARSTEN AND EDITH BROOKS 302

A Mechanical Determination of Correlation Coefficients and Standard Deviations . . . . JOHN R. PLATT 311

The Evidence for Periodicity in Short Time Series . TRUMAN L. KELLEY 319

Correlation Concepts and the Doolittle Method . . DUDLEY J. COWDEN 327

Overestimation of Mean Squares by the Method of Expected Numbers . . . . R. E. COMSTOCK 335

Utility of Statistical Method in Aerodynamics . HERBERT G. SMITH 341

On Measures of Dispersion for a Finite Distribution . . . . ALBERT O. HIRSCHMAN 346

On Some Census Aids to Sampling . . . . MORRIS H. HANSEN AND W. EDWARDS DEMING 353

The Social Insurance Movement . . . . . . R. CLYDE WHITE 358

Horace Secrist, 1881–1943 . . . . F. S. DEIBLER 365

Committee on Nominations 366

BOOK REVIEWS

ALTMAN, GEORGE T. *Federal Tax Course.* James D. Paris 373

BARGER, HAROLD. *Outlay and Income in the United States, 1921–1938.* Everett E. Hagen 368

CHASE, STUART. *The Road We Are Traveling.* Forrest H. Kirkpatrick . . 377

CLARK, CARRIE PATTON, see CLARK, FRED E.

CLARK, FRED E., AND CLARK, CARRIE PATTON. *Principles of Marketing.* Howard T. Hovde 376

FURNESS, J. W., see LEITH, C. K.

GRAS, N. S. B. *Harvard Cooperative Society, Past and Present.* Colston Warne 379

Continued on next page

HARTKEMEIER, HARRY PELLE. *An Introduction to Managerial Business Statistics.* Donald R. G. Cowan 382

KATONA, GEORGE. *War without Inflation.* A. Smithies 367

LEITH, C. K., FURNESS, J. W., AND LEWIS, CLEONA. *World Minerals and World Peace.* Wilbert G. Fritz 375

LEWIS, CLEONA, see LEITH, C. K.

MACHLUP, FRITZ. *International Trade and the National Income Multiplier.* Paul A. Samuelson 369

MIGONE, RAUL C. (Editor) *Inter-American Statistical Yearbook, 1942.* Myron E. Andrews 374

NIENSTAEDT, L. R. *Economic Equilibrium, Employment and Natural Resources.* Elmer C. Bratt. 371

ROBISON, SOPHIA M. (Editor) *Jewish Population Studies.* Georg Wolff . 381

VON MERING, OTTO. *The Shifting and Incidence of Taxation.* Carl S. Shoup. 372

PUBLICATIONS RECEIVED. 384

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 223 SEPTEMBER, 1943 *Volume 38*

OBSERVATIONS ON THE COST OF LIVING INDEX OF THE BUREAU OF LABOR STATISTICS*

BY LAZARE TEPER
Director of Research, International Ladies' Garment Workers' Union

THE COST OF LIVING index[1] of the Bureau of Labor Statistics hopes to measure "changes from time to time in prices to the ultimate consumer of goods purchased by a representative group of wage earners and lower salaried workers in the larger cities of the country" whose average incomes in the 1934–36 period amounted to $1524. Data are collected for 34 cities. The cost of living index for each of the cities is a combination of index numbers for the following groups representing workers' spending: food; clothing; rent; fuel, electricity and ice; housefurnishings; and miscellaneous items, which include among others transportation, medical care, accident and health insurance, household operation, recreation, personal care and gifts and contributions. When prices are secured in a particular city for a specific commodity quotations for each individual service or commodity are averaged. The Bureau then calculates the percentage changes in these average prices from those in the previous period, using an identical sample of items and outlets, and applies them to the estimated costs of each item in the previous period to obtain costs in the current period. A grand total and totals by budgetary groups are computed. These totals are related to the totals in the base period to obtain the indexes. Under this procedure the Bureau applies price ratio to basic

---

EDITOR'S NOTE. Mr. Teper's article is the first of two or three documents on the Bureau of Labor Statistics index of the cost of living which will be run in this and following issues of our JOURNAL. At the time of going to press a Committee of the American Statistical Association is actively reviewing the cost of living index. This Committee was requested by the Bureau of Labor Statistics and the Secretary of Labor. It consists of Professor Frederick C. Mills, of Columbia University, *Chairman;* E. Wight Bakke, Yale University; Samuel S. Stratton, Middlebury College, Vermont; Theodore W. Shults (Miss Margaret Reid, alternate), both of Iowa State College; and Reavis Cox, University of Pennsylvania. It is hoped that its report will appear in the December issue. In some succeeding issue the Bureau of Labor Statistics may wish to publish further statements regarding the index.

* This paper was presented at the Conference of Research Directors of National and International Unions in Washington, D. C., on June 23, 1943.

The writer wishes to acknowledge the assistance rendered by Nathan Weinberg and Sheila Helmann of the ILGWU Research Department in the preparation of the paper.

[1] U. S. Bureau of Labor Statistics, Bulletin No. 694: *Handbook of Labor Statistics*, 1941 edition, I, p. 83.

expenditures by wage earners and lower-salaried clerical workers to secure cost aggregates for each pricing period.[2]

The "national" cost of living index is derived from the index numbers of the individual cities. Each of the component indices is weighted by the relative importance of the population of the particular city, augmented frequently by the population of not too far distant communities of 50,000 and more in population.

Measurement of changes in living costs is important to the wage earner and his family at all times. Low income-bracket families feel keenly fluctuations in the purchasing power of their wage dollar. In periods like the present, when the pressure of rising prices makes itself felt on the one hand, and when the ability to secure wage adjustments is limited by governmental edict on the other, interest in cost of living figures and their accuracy is naturally uppermost in many minds. This interest is spurred on by the natural bias of the individual who thinks of rising prices in terms of the few commodities which have risen most in price. This mental weight given to those commodities which lead the price parade frequently obscures the real picture and makes for a not too well justified criticism of governmental figures. However, criticisms of the present Bureau of Labor Statistics cost of living index are not restricted to the biased. Labor economists, business publications, and government officials recognize today that the cost of living figures leave much to be desired.[3] In all fairness it should be noted that it does indicate, even if occasionally and in a weak voice, that "cost of living indexes have certain limitations which should be remembered by those who use them,"[4] as well as that "the index is an approximate measure."[5]

Considerable amount of criticism arises through the fault of the Bureau itself. It is much too eager at times to defend its product. Thus despite all the difficulties which are inherent in securing black market prices, the Bureau is emphatic in stating that its figures "are based on actual selling prices, regardless of OPA ceilings."[6] Yet, prevailing opinion does not credit the Bureau's contention. Professor Irving Fisher, discussing recent increases in the cost of living, states that "probably the rises were actually greater than these records show, since the records can take no account of the 'black markets'—the bootlegging of goods outside of the regular channels of trade."[7] A business periodical honestly remarks:

---

[2] Robert A. Sayre, "Cost of Living Indexes," in this JOURNAL, June 1941, p. 193. See also U.S Bureau of Labor Statistics, Bulletin No. 699: *Changes in Cost of Living in Large Cities in the United States, 1913–41*, p. 35.

[3] To cite but one example of interdepartmental criticism, we may refer to the U. S. Department of Commerce, *Survey of Current Business*, February 1943, p. 2, which speaks thus of the Canadian as well as the United States indexes: "Both the indexes probably understate the true rise in living costs because of quality deterioration, illegal price advances, and changes in consumption patterns."

[4] U. S. Bureau of Labor Statistics, Bulletin No. 699, p. 8.

[5] "What is the Cost-of-Living Index?" in *Monthly Labor Review*, August 1942, p. 273.

[6] U. S. Bureau of Labor Statistics, *The Cost of Living and Retail Costs of Food*, March 15, 1943 (mimeographed release). A somewhat different statement appears in *ibid.*, January 15, 1943: "The Bureau of Labor Statistics' cost of living index reflects actual prices in retail stores where families with moderate incomes trade. Black market operations or sales to customers who pay bonuses for service cannot, however, be measured." The writer does not imply that the Bureau does not secure a certain number of over-the-ceiling price quotations.

> These illegal price gains, of course, will not show up in official indexes of the cost of living. Many items are priced at the carefully complying big stores, and, in any case, the obvious federal agent will hardly be quoted above-ceiling prices by even the least sharp-eyed seller.[8]

A government official categorically states that "Black markets and the inability to obtain supplies at any price are not, of course, measured by price statistics," even though he notes that "the food stores reporting to the Bureau of Labor Statistics actually report their own violations in a surprising number of instances"[9] either through confidence in the Bureau or because of their ignorance of regulations. It is obvious that nothing short of actual purchase by field agents will provide that Bureau with information as to the "actual selling prices."

Nor does the Bureau, in discussions of its cost of living index, adequately point out the real nature of the problem involved in pricing identical commodities at different pricing periods.[10] Stress in the Bureau's publications is laid on the fact that with present day techniques a "more consistent price reporting is possible."[11] This, of course, is an accurate observation, but it begs the question. Granting that progress has been made, how much further does the Bureau have to go to insure adequate comparability of goods to be priced? The existing specifications, although more detailed than those of several years ago, continue to be couched frequently in vague, suggestive terms, rather than in terms which permit comparability controls.[12] The statement that

---

[7] Irving Fisher, "Inflation; Its Chief Causes and Remedies," in *Union Label Feature Service*, Release no. 31, January 9, 1943, p. 3.

[8] *Business Week*, December 19, 1942, p. 13.

[9] Don D. Humphrey, "The Effectiveness of Price Control," in U. S. Department of Commerce, *Survey of Current Business*, February 1943, p. 24.

[10] In many fields competitive rivalry centers largely or exclusively upon non price elements. "The nature of the commodity may be such as to make price comparisons difficult and emphasis upon non-price factors logical. When this tendency exists, it may be reinforced by the development of conventional patterns of business practice." (Temporary National Economic Committee, Monograph No. 1: *Price Behavior and Business Policy*, p. 70). Under such conditions considerable danger exists that the continued existence of a particular price line may be construed as no change in price, and that price changes will only be recognized when there is a shift to new price lines. To the extent that such approach may exist, the indexes for the individual commodities merely establish the fact that a commodity of the particular general class can be obtained at a particular price. As such the index fails of course to fulfill the theoretical requirements of a cost of living measure.

[11] U. S. Bureau of Labor Statistics, Bulletin No. 699, p. 6.

[12] See for example *ibid.*, p. 7, for a specification of a men's medium quality wool suit. As any person in the trade could point out, considerable variations in the quality of the suit are possible within the framework of this definition.

> since 1935, the Retail Price Division of the Bureau of Labor Statistics has been using quite precise commodity specifications in pricing at retail, in order to be assured that prices at two successive periods refer to the same commodities—the same not only in size and appearance, but in quality as well[13]

clearly overstates the case. The memory of the field agents and of the retail store personnel[14] must of necessity constitute the major control as to what a particular article was like at the time of the last visit of the agent as against the article priced during the current visit.[15] No experimental controls have been attempted by the Bureau to test the degree of possible error in judging comparable qualities. Even if the "scientific testing of consumers' goods for durability and for efficiency in performing the services for which they are intended is still in its infancy,"[16] the fact remains that

> Performance tests have been developed by the Federal government in order to improve its purchasing procedures, and in order to administer the tariff or acts intended to protect the consumer, or to improve trade practices. Other performance tests have been worked out by industrial engineers working for manufacturers or distributors. In some cases these industry tests have been reviewed by the scientific societies concerned and the American Standards Association, and have become standard practice. . . . [17]

It would seem that it is a function of the Bureau to develop, either on its own behalf, or jointly with the scientific organizations, available information on performance. Such information could then be utilized first for determining the degree of accuracy in the field agents' estimates of comparability. Once the extent of error is established, this information could be used to introduce a "corrective factor" into the index which would account, in a more precise fashion than at the present, for the "hidden price changes."[18] At present, however, the initiative which the Bureau could exercise is dissipated, while the problem of accounting for quality changes remains unsolved.

---

[13] Faith M. Williams, Frances R. Rice and Emil D. Schell, "Cost of Living Indexes in Wartime," in this JOURNAL, December 1942, p. 419.

[14] It is conceded presumably that there exists an "inadequacy of retailers' knowledge of their own merchandise" (*Ibid.*).

[15] An interesting side light is thrown on this problem by the affidavits prepared by the technical personnel of the OPA Standards Division in connection with its investigation of price violations by the Hecht Co., Washington, D. C. These searching investigations revealed that, despite the admitted effort made by this company to comply with the OPA price ceiling, the store buyers were unable to locate, without considerable error, garments of comparable quality to those sold during the base period. (Affidavits of Raymond J. Singer and Robert E. Vanbrunt, in the files of the District Court of the United States for the District of Columbia, Civil Action File No. 17995: Leon Henderson v. The Hecht Co.)

[16] Williams, Rice and Schell, *supra*. p. 420.

[17] *Ibid.*

[18] To implement such a program, agents of the Bureau must purchase, at different pricing periods, a certain number of articles on which they normally obtain quotations. This program could be combined with purchases necessary to determine the extent to which black market prices prevail.

A certain amount of misunderstanding arises from the fact that, by and large, except for the description of the general make-up of the index, little is known, even among those interested in the subject, about the actual handling of the data once it is collected. A case in question is well illustrated by the Bureau's study of its new index, published as its Bulletin No. 699. The advent of this study was made known in a previously published article in the *Monthly Labor Review* which stated:

> A detailed consideration of all the steps involved in computing the new index is outside the scope of the present article. A bulletin giving, in addition to a *detailed account* (my italics—L. T.) of the various procedures employed, a list of the weights for each city is now in preparation.[19]

This promise however was but partly kept. Bulletin No. 699 does give the population weights used for combining index numbers for the individual cities into composite indexes for the United States. But it does not give a "detailed account" of all the procedures employed. For one thing the serious problem of securing price quotations is treated most generally.[20] Yet it is a neglected subject in the literature of statistics,[21] certainly if compared to the amount of space in statistical literature devoted to the lesser question of the index number formulae.[22] One could certainly wish the Bureau would devote considerably more space to the public discussion of this problem. Free and easy access to specifications used by the Bureau in collecting retail prices is not enough. Instructions issued to field agents (other than those which are purely administrative in character) should be regularly released. All changes in specifications which find their way into the monthly tabulations should be made known. Special problems encountered in the course of collecting data should be given. Similarly it would be desirable for the Bureau to release regularly its instructions to the personnel entrusted with editing and tabulating information secured in the field. It is hardly satisfactory to be told in general terms that some adjustments were made in the index.[23]

[19] "U. S. Bureau of Labor Statistics' New Index of Cost of Living," in *Monthly Labor Review*, August 1940, p. 369.

[20] The entire discussion of this subject is compressed into less than 3 pages.

[21] This does not mean that the subject received no attention whatsoever. Wesley C. Mitchell's pioneer work, "The Making and Using of Index Numbers" devotes several pages to the topic (U. S. Bureau of Labor Statistics, Bulletin No. 284, *Index Numbers of Wholesale Prices in the United States and Foreign Countries*, pp. 25 ff.). The subject was forcefully brought to the fore by Willard L. Thorp (see his "Price Theories and Market Realities" in *American Economic Review*, March 1936, supplement, pp. 14 ff.). An interesting discussion of the problem is also found in the U. S. National Resources Committee, *The Structure of the American Economy*, I, 173 ff. See also Williams, Rice and Schell, *supra*.

[22] "The discussion concerning index number technique has turned mainly on the selection of data and the algebraic characteristics of various formulas, with insufficient attention being given to the fundamental nature of the phenomena dealt with." (Wirth F. Ferger, "Distinctive Concepts of Price and Purchasing-power Index Numbers," in this JOURNAL, June 1936, p. 258).

[23] An example of such generalities may be found in the U. S. Bureau of Labor Statistics, "Descrip-

If such information were generally available it could go a long way toward clearing up misunderstandings between the Bureau and the users of its figures. Such information could readily be incorporated into special technical bulletins, to be released at monthly intervals, to interested persons and groups. These technical bulletins could also serve another function—a discussion of the theoretical problems concerning the significance of the index, and their reconciliation with the practical limitations in the index construction. In theory the cost or price of living may be defined as "the ratio between incomes securing the same standard of living, or, in other words, procuring the same sum total of economic satisfaction, in two situations which differ only in respect of prices."[24] Even if the above relationship holds true between two consecutive pricing periods, what is the significance of the index when radical changes in the actual level of living which is priced take place over a period of time? It is usually agreed that when "the links are chained together, precise comparability is lost, and the meaning of the index cannot be stated in simple language."[25] What biases are introduced by chaining link relatives to a single index? Warren Persons notes that

> It has frequently been observed, in actual practice, not only that chain indexes with changing weights differ from corresponding fixed base indexes, but also that the divergence of the two indexes tends to persist and even to become irregularly greater and greater as the number of links in the chain increases.[26]

The problem has been recognized by the Bureau of Labor Statistics in connection with its index of wholesale prices,[27] and also in connection with the data on employment and payrolls.[28] What recognition has been given to this factor in connection with the cost of living index?

---

tion of Changes in the B.L.S. Cost of Living Index as of March 15, 1943" (mimeographed release). This release states that "adjustments were made for changes in the volume of food sold through chains and independent stores and supermarkets." What was the nature of these adjustments? On what information did the Bureau rely in making changes? What provisions if any were made for any future shifts as between retail outlets? Similarly it is important to know when certain cases of substitutions are introduced not by a linking method but as price changes (U. S. Bureau of Labor Statistics, Bulletin No. 699, p. 7).

[24] International Labour Office, Studies and Reports, Series N (Statistics), No. 20, *International Comparisons of Cost of Living: A Study of Certain Problems Connected with the Making of Index Numbers of Food Costs and of Rents*, p. 74.

[25] Frederick E Croxton and Dudley J. Cowden, *Applied General Statistics*, p. 621.

[26] Warren Milton Persons, *The Construction of Index Numbers*, p. 65.

[27] Jesse M. Cutts and Samuel J. Dennis, "Revised Method of Calculation of the Wholesale Price Index of the United States Bureau of Labor Statistics," in this JOURNAL, December 1937, pp. 665 f.

[28] "The weakness of chain indexes is the possibility of errors of a progressive and cumulative character. With a chain type of index an error in calculating the percentage of change in any one month not only affects the index for that particular month but is carried to all subsequent items of the series." (U. S. Bureau of Labor Statistics Bulletin No. 610: *Revised Indexes of Factory Employment and Pay Rolls, 1919 to 1933*, p. 11.)

The Bureau admittedly is making a number of adjustments in the index which affect its make-up.[29] Quantity weights are changed to allow for changes in the supply situation. To the extent that the needs of war economy bring about a deterioration in living standards,[30] budgets now priced by the Bureau (even if we ignore subtitutions for those goods and services which are no longer available) are representative of a lower level of living. Do we on that basis obtain valid comparisons of the changes in the *cost* of living with the pre-war period? Do the substitutions of commodities made today affect the index to the same extent as substitutions in a period of normalcy, when changes in the types of available goods and in consumption habits are a slow and gradual process? Should the Bureau change consumption weights for those items that continue to be available in civilian markets?[31] Should not the Bureau preserve, to the extent that it is possible, the level of consumption priced in pre-war days and allow in the budgets which it prices equivalent expenditures (at the time substitutions are made) in terms of commodities and services which replace the ones no longer obtainable?[32] On the other hand, should a different view be taken, can the Bureau contend that its index measures only changes in the cost of living and not of the plane of living as well?

The chaotic situation in retail markets during war time brings another problem to the fore. Are individual price samples collected by the Bureau reliable indicators of price movements in particular cities? When it comes to commodities and services other than food or rents, the Bureau is satisfied as a rule to secure four quotations, except in the city of New York where five quotations are secured.[33] This is a fact not well enough known.[34]

[29] Williams, Rice and Schell, *supra;* U. S. Bureau of Labor Statistics, *Description of Changes in the B.L.S. Cost of Living Index as of March 15, 1943.* Also "Bureau of Labor Statistics Cost of Living Index in Wartime," *Monthly Labor Review*, July, 1943.

[30] "As the production machinery of the country is converted to war purposes, the goods and services produced for consumers are undergoing more drastic changes in the course of a single year than occurred in the entire period from 1919 to 1935—but in the opposite direction." (Williams, Rice and Schell, *supra*, p. 416.)

[31] A question well may be asked regarding the treatment of individual price series at the time the cheaper grades of a particular commodity disappear under the stress of war economy and only more expensive grades remain—how are such situations to be treated? It is the practice of the Bureau to treat such situations not by linking them with the older series but as price changes. (U. S. Bureau of Labor Statistics Bulletin No. 699, p. 7.) This method is not inconsistent with the point made in the text which applies primarily to situations when goods which disappeared are replaced with quotations of commodities in different classifications. Forced upgrading of the consumer is definitely in the same class with situations when under the stress of necessity the worker is forced to trade in more expensive stores.

[32] A somewhat similar principle has been recognised in the construction of the index of wholesale prices. When a particular series is replaced with another, an adjustment of the "quantity multiplier to take account of the substitution" is "made on the basis of the typical relationship between the two series" (Cutts and Dennis, *supra*, p. 669).

[33] U. S. Bureau of Labor Statistics, *Retail Prices: Instructions for Collecting Retail Prices, Commodities and Services (Excluding Rents)*, September 1940 (mimeographed), paragraph 138, p. 16, gives the following instructions to field agents: "Secure four quotations for each item on the schedules in every

It is questionable whether a sampling of 4 or 5 similar (or identical) commodities yields a representative value. Let us consider an example. According to the Census of Distribution for 1939, dresses were retailed by 567 stores in New York City divided as follows:[35]

Type of store	Number of stores reporting sales of dresses	Dress dollar volume of reporting stores
Department stores	25	$11,373,000
Family clothing stores	35	$ 8,959,000
Women's ready to wear stores	507	$58,217,000

The above figures do not represent all stores selling dresses in the city of New York as the 1939 census records the existence of 2818 stores (in the above listed classifications) doing a total dollar volume of $502,099,000 as against 632 stores doing a total volume of $435,165,000 which have reported product breakdowns. While it is undoubtedly true that not all of these stores catered to a clientele of wage earners and lower salaried workers, it is nevertheless obvious that quotations from 5 stores are hardly representative of prices in the New York retail dress market. The sample is so small that it is doubtful whether any test of significance could be devised which could be applied to the

---

city except New York where five are to be secured." For list of commodities to which these instructions apply see *ibid.* pp. 18 f. Undoubtedly in some instances more than one article in the same general class may be priced. "For a large proportion of the items included in the index more than 1 quality is priced; in the case of the more important items, as many as 4 in a given city." (U. S. Bureau of Labor Statistics Bulletin No. 699, p. 15.)

[34] U. S. Bureau of Labor Statistics Bulletin No. 699, although presumably a complete discussion of the index and of its construction does not mention it. It does however mention the number of outlets reporting retail food prices for June 1941 (*ibid.*, p. 88). According to another Bureau's publication, the number of firms reporting food prices varies from 10 in the smallest cities to 50 in the largest "roughly proportional to the population (U. S. Bureau of Labor Statistics, Bulletin No. 635: *Retail Prices of Food, 1923–36*, p. 10). Nor does this study indicate the size of the rent sample other than stating that "Advantage was taken of the Real Property Inventory and of local studies of housing to secure a sample which would be representative of housing conditions in the cities covered." (U. S. Bureau of Labor Statistics, Bulletin No. 699, p. 20.)

Miss Faith M. Williams writes on July 16, 1943, as follows: "The number of outlets from which prices are obtained in each city was inadvertently omitted from Bulletin 699. It had been published previously, however, in the August, 1936, February, 1938, and November, 1938, pamphlets on Retail Prices and in a mimeographed publication on retail prices of hosiery by areas (June, 1940). In response to requests for average prices of commodities other than food, the Bureau always explains that its reason for not releasing such figures is that only four outlets are covered in each city except New York where five are covered. In considering the number of outlets covered for each commodity priced, one must take account of the fact that for the clothing and housefurnishings included in the index, more than one quality and for some goods as many as four qualities are priced in each outlet. There is undoubtedly an error of estimate in average price changes based on four quotations, but there is no reason to assume that there is a bias in the errors of estimate. The clothing index for each city for each date is based on 421 quotations; the housefurnishings index on 129 quotations. The Bureau does publish per cent changes in costs of individual items for all cities combined and each one of these is based on 137 price quotations for each date." Miss Williams' assumption that there is no bias in the errors of estimate, i.e., that the errors of estimate remain constant over a period of time, is unduly optimistic.

[35] U. S. Census Bureau, *1940 Census of Business*, I (2), 234, 310, 312.

data so collected. The presence of pronounced dispersion in the data so collected may further weaken the validity of the average quotation. It is interesting to note in this connection that instructions to the Bureau's field agents make the following comment on the presence of dispersion in the price quotations secured in a particular city:

> Dispersions. Prices on an item conforming to the same specification will usually show some dispersion. However a range of prices which exceeds 50 per cent from lowest to highest is considered unusual and must be explained on the tally.[36]

It is therefore imperative for the Bureau to increase the size of the samples of many of the commodities it prices. For only in such a manner can it hope to eradicate the errors of inadequate sampling.

Another observation pertaining to the securing of price quotations made be made at this point. The Bureau customarily obtains regular prices from its contacts in the retail field. Sale prices, i.e., markdown prices, according to the Bureau's instructions to its field agents may be taken under the following conditions:

> The period of sale must be two weeks or longer and must include the date of the pricing period.
>
> The sale must not be one for the purpose of discontinuing the item permanently. This does not refer to seasonal clearance sales.[37]

Similar instructions are issued in connection with the collection of retail prices of food:

> Only the "regular" price is to be obtained for individual items. This excludes clearance, distress and bankruptcy prices, and also special offers. "Special sale" prices may be accepted where the "special" prevails for several days.[38]

In normal times the wage earner and his family frequently resort to buying "specials" at sales. To the extent that sales, whether of one day's duration or longer, continue from year to year and are responsible for an approximately constant amount of the consumer dollar, their incomplete recognition by the cost of living index would not materially influence its trend. However, in times like the present, the relative importance of special sales has materially declined. While the 1942 figures have not been released at the time of this writing by the National Retail Dry Goods Association, an analysis of the relative importance of markdown in department stores between 1935 and 1941 shows definitely that they are on a decline.

---

[36] U. S. Bureau of Labor Statistics, *Retail Prices: Instructions for Collecting Retail Prices, Commodities and Services (Excluding Rents)*, September 1940, paragraph 155, p. 17.

[37] *Ibid.*, paragraphs 141 and 142, p. 15.

[38] U. S. Bureau of Labor Statistics, *Retail Prices: Instructions for Collecting Retail Prices, Food*, March 1942 (mimeographed), paragraph 18, p. 2.

Markdowns[39] as Per Cent of Retail Sales, Department Stores with Annual Sales volume over $500,000, 1935–1941

Year	MARKDOWN PERCENTAGE	
	Main Store	Basement
1935	7.1	6.9
1936	6.6	6.3
1937	7.0	6.5
1938	7.3	6.6
1939	6.9	6.2
1940	6.5	6.1
1941	5.4	5.3

When the volume of special sales undergoes a significant variation, it is desirable to record their influence in the cost of living index. While it is manifestly impossible to secure sales quotations, particularly if sales are prone to be of short duration, and while it is difficult to obtain the relative volume of business done at regular prices as distinguished from the sales prices, it may be possible to work out a series of correction factors which, though not obtainable on a current basis, could nevertheless be periodically incorporated in the index for the purpose of correction of price data. If the relative importance of markdowns is known for a class of commodities, it may be used for the purpose of constructing such an adjustment factor.[40]

An adjustment factor is probably needed in connection with the changes in the Bureau's practice in collecting data on rents. In the past rental quotations were obtained from rental agencies. However, beginning with September 1942, rent quotations in 13 cities were obtained from tenants instead.[41] (This practice has now been extended to all cities.) If discrepancies were found between quotations secured from rental agencies and those of tenants as a result of war conditions, allowance should have been introduced in the cost of living index.[42]

The cost of living index of the Bureau of Labor Statistics is supposed to measure the cost of goods and services purchased by wage earners and lower-salaried workers. Yet one of the services paid for by them is but partially represented in the index—the cost of maintaining the services and functions of government. Sales taxes are added to the cost of the commodities on which they are imposed. Automobile and other taxes are specifically included. Property taxes are indirectly reflected in rentals. However, income taxes and social security taxes are not included in the index.

---

[39] National Retail Dry Goods Association, Controllers' Congress, *Departmental Merchandising and Operating Results of Department Stores and Specialty Stores* (for the years 1935 through 1941).

[40] If in a particular year the percentage of mark-downs to the volume of sales (S) amounted to m, then, for each of the years which we wish to study $S=pq$, where p is the average price actually received by the merchant for his merchandise, and q is the quantity sold. If there were no "special" sales responsible for markdowns then the "value" of sales (V) would have been $V=S(1+m)=q[p(1+m)]$. If it is the practice of field agents to get the quotations of "regular" prices, i.e. $p(1+m)$, then the actual average prices obtained from the store's customers can be computed, when m is known.

[41] Williams, Rice and Schell, *supra*, p. 421.

[42] Miss Faith M. Williams writes on July 16, 1943, as follows: "Discrepancies were found between percentage changes shown by quotations obtained from rental management agencies and in the rents as reported in the tenant survey in four cities. The following figures show the indexes as originally published and as revised. Differences between the averages from the rental management surveys and the tenant

The reasoning for the exclusion of income taxes is not altogether clear. The *Handbook of Labor Statistics* suggests that "income taxes paid have . . . been omitted as they have heretofore applied to a very small proportion of the groups whose living costs the indexes attempt to measure."[43] A recent article by three officials of the Bureau who are concerned with the compilation of the index, however, rationalized the exclusion on the ground that "income taxes . . . are considered as deductions from income and not as current expenditure."[44] The Bureau's bulletin intended as a reference book for persons using the cost of living indexes merely states, "Income taxes paid have also been omitted."[45]

Prior to the present war, for a number of years, income taxes in the low income brackets were insignificant in amount, if any tax was paid at all. Furthermore there was little variation in the tax rates. They could therefore be disregarded in the computation of cost of living indexes, as they would not have affected the course of the index in the slightest degree. Since those days, however, the tax picture has undergone a radical change. Even persons in the lowest income brackets must pay something toward governmental services. Little justification exists

---

surveys were not significant in the other cities covered."

(1935–39 = 100)	Rental management survey	Tenant survey
Cleveland		
June 1942	112.0	115.7
July 1942	112.0	115.5
August 1942	112.1	115.3
Detroit		
June 1942	112.5	115.1
July 1942	111.0	114.9
August 1942	110.8	114.7
Manchester		
September 1942	107.8	107.6
Portland, Oregon		
September 1942	115.1	115.6

[43] U. S. Bureau of Labor Statistics, Bulletin No. 694, I, p. 84.

[44] Williams, Rice and Schell, *supra*, p. 422.

[45] U. S. Bureau of Labor Statistics Bulletin No. 699, p. 8.

under such conditions for treating income taxes differently from direct taxes used for financing similar or identical services.

The inclusion of income taxes in the cost of living index is not a new idea. The Swedish cost of living index includes income taxation as one of its components.[46] The International Labour Office suggested definite procedures on the inclusion of income taxes in the index.[47] Representatives of the Bureau of Labor Statistics agree that "the amount of the tax to moderate-income families can, of course, be readily computed."[48] Little reason therefore remains for continuing to exclude the cost of income taxation from the cost of living index.

Social security taxes present a somewhat different problem. Their exclusion from the index is justified on the ground that they "have been treated as savings."[49] However, social security taxes represent savings only in part.

It is interesting in connection with this point, to examine the reasoning of the Bureau of Labor Statistics justifying the exclusion of life-insurance premiums from the computation of the cost of living figures and for treating them as savings:

> It is recognized that most insurance-policy premiums include payments for several elements, only one of which is truly savings. The first is for actual cost of life-insurance protection during the year in question. This would amount to the cost on an actuarial basis of term insurance for 1 year at the actual age of the insured. Such cost is properly current family expenditure for insurance protection for the year. Another element is the part of the premium which goes toward operating costs of the insurance company. This element is especially large in the case of industrial insurance, which covers the expense of making weekly collections. This element is also not properly savings, but merely a form of current family expenditure. Any amounts included in the premium payments in excess of these two items, which accumulate in the form of net cashable value of the policy, are truly savings. To the extent that policies are allowed to lapse under terms which mean loss of payments previously made, even such payments can only doubtfully be classed as savings. . . . It is . . . impossible to estimate how much of the amount paid in life-insurance premiums represents savings and how much was paid for insurance protection during the year or other services of the insurance company. The entire amount of such payments has therefore been treated as a disposition of funds for items other than current family expenditure, an increase in assets, and hence as savings. . . .[50]

---

[46] Williams, Rice and Schell, *supra*, p. 422.

[47] International Labour Office, Studies and Reports, Series N (Statistics), No. 6: *Methods of Compiling Cost of Living Index Numbers*, Report prepared for the Second International Conference of Labour Statisticians, pp. 45 ff.

[48] Williams, Rice and Schell, *supra*, p. 422.

[49] U. S. Bureau of Labor Statistics Bulletin No. 699, p. 8.

[50] U. S. Bureau of Labor Statistics Bulletin No. 638: *Money Disbursements of Wage Earners and Clerical Workers, 1934–36*, Summary Volume by Faith M. Williams and Alice C. Hanson, p. 179.

It is evident from the above that the only reason for the non-inclusion of the life-insurance premiums is a technical one: inability to separate those elements in life-insurance premiums which represent the current, actual cost of insurance to the policy-holder from those which represent savings. Were such a separation possible an allowance would have been made.

It is obvious that this logic holds true for other forms of insurance as well. Thus we find that the Bureau's consumer expenditure studies treat automobile accident and health insurance as expenses.[51] These items are incorporated in the cost of living index.[52]

The analogy between the treatment of insurance taken out with private carriers and social security contributions is perfect. Unemployment insurance taxes, in those states where the worker contributes, are clearly comparable with term insurance. The payment of taxes buys for the worker certain limited insurance protection. There is no paid up value. Unless he collects benefits within a benefit year, he has no accrued rights. It thus becomes a form of current family expenditure.

On the other hand it can hardly be argued that old age and survivor benefits come within the same class. While there is no paid in value, these contributions definitely can be characterized as savings.

Two omissions from the index may be mentioned in brief. One of these is represented by restaurant prices. The Bureau had planned to incorporate meals eaten away from home into the index during the fiscal year 1942.[53] To date this reform has not been carried out. The other omission is the recognition that wage-earners who own their own homes have a somewhat different problem from their rent-paying neighbors. With the general rise in prices and changes in tax rates the cost of home-ownership may follow different trends from rentals. The problem certainly deserves investigation.

In order to secure a national index of the cost of living the Bureau weighs indexes for the individual cities by the population of the given metropolitan area and that of other cities in the same region and size class.[54] To what extent does an index for a specific city represent the cost of living changes of a larger area comprised of several towns? A check with special studies of cost of living changes made by the Bureau of Labor Statistics shows that it is not always representative. Comparing, for example, the cities of Boston and Bridgeport we find that between October, 1939, and October, 1942, the following changes in the cost of living took place:

---

[51] *Ibid.*, pp. 138 ff. and pp. 158 f.

[52] U. S. Bureau of Labor Statistics Bulletin No. 699, p. 23.

[53] U. S. Bureau of Labor Statistics Bulletin No. 699, p. 16.

[54] *Ibid.*, p. 39. Recently the Bureau modernized its city weights to allow for wartime changes in population (U. S. Bureau of Labor Statistics, *Description of Changes in the B.L.S. Cost of Living Index as of March 15, 1943*).

	Boston[55] (Per cent)	*Bridgeport* (Per cent)
All items	19.3	22.1
Food	34.7	33.9
Clothing	22.7	29.6
Rent	4.9	6.0
Fuel, electricity and ice	24.1	13.5
House furnishing	18.7	22.3
Miscellaneous	10.0	15.8

While a single example does not make the rule, it indicates the desirability of experimentation and the inclusion of a larger number of cities, at least periodically, for the purpose of checking possible deviations of the national index.

NOTES ON MR. TEPER'S OBSERVATIONS

BY ARYNESS JOY WICKENS
Chief, Prices and Cost of Living Branch
AND FAITH M. WILLIAMS
Chief, Cost of Living Division, Bureau of Labor Statistics

MR. TEPER's article on "Observations on the Cost of Living Index of the Bureau of Labor Statistics" expresses ably some of the doubts existing in a number of quarters concerning the accuracy of the index, and raises a number of important technical questions. Most of these will be discussed by the forthcoming report of the Committee of the American Statistical Association.

The Bureau of Labor Statistics would prefer to let the report of the Committee speak for itself on most of these points. There are, however, a number of points which are worthy of note at this time.

(a) The Bureau of Labor Statistics has not attempted to obtain prices charged on "black markets" outside the regular channels of trade. The index *does*, however, contain many prices which are above Office of Price Administration ceiling levels. For certain cuts of meat, on a recent occasion just after dollars and cents ceilings had been established, 25 per cent of the prices reported were above ceilings; for rents, a recent comparison shows about 4 per cent above ceilings in certain cities. Moreover, there have been repeated instances of reported prices which increased month after month in the same store, although they were under the General Maximum Price Regulation and should not have advanced.

---

[55] October, 1939, figures (other than food) for Boston were estimated by interpolating between September and December, 1939, indexes.

It is not possible to say what percentage of the total number of quotations in the index is above ceiling levels because the great variety of ceiling regulations in different areas and for different products makes comparisons difficult.

Nor is it contended by the Bureau of Labor Statistics that *all* across-the-counter prices above ceilings (as distinguished from bootleg or "black market" sales) are reported. Certain special tests made for the use of the American Statistical Association Committee indicate, however, a close correspondence of actual sales prices charged customers with prices given to the representatives of the Bureau.

(b) Recently, fewer stores have held "white sales," and the price declines have been less than in the pre-war period. This is also reflected in the cost of living index. No figures are available to indicate the extent to which wage earners and lower salaried groups have in the past made use of special sales, or the extent to which they are using them now. It is therefore impossible for the Bureau to devise correction factors for the cost of living index which would account for the wartime reduction in the number of special sales. Where the whole market goes down for an extended period of a month or more, as in the case of the "white sales," and the practice is accepted throughout the trade, these lower prices are reflected in the Bureau's monthly index, but without added weight. They are treated, in effect, like any other seasonal price decline.

(c) With regard to measuring changes in the serviceability of consumer goods at different price levels by laboratory tests, the Bureau has long been interested in the problem of quality change, but has found no satisfactory way to deal with it. Mr. Teper quotes part of a paragraph from an article which appeared in this Journal in December 1942. That paragraph went on to say that the Bureau had explored at some length in the summer and fall of that year the available test material on changes in the serviceability of consumer goods due to the use of substitute materials necessitated by the war situation, and had found no summary data which could be applied as quantitative factors to take statistical account of hidden price changes. Mr. Teper suggests that the Bureau should utilize available material on the performance of consumers' goods as a corrective factor in the cost of living index. Such a procedure would not only be exceedingly difficult, but would require a large appropriation for the purchase and testing of the goods priced at different periods, and extensive research as to methods of summarizing measures of durability, serviceability, etc., as for example, measures of tensile strength, resistance to abrasion, air permeability, and heat conductivity for overcoating. It is very doubtful, in our opinion, if this proposal is practicable.

In evaluating the desirability of carrying out Mr. Teper's suggestion, even if it were practicable to do so over a long period of time, it is important to realize that in periods of technical development when the quality of consumers' goods has been improving the Bureau has faced a similar problem of quality change, but in the other direction. In the pre-war period the quality of the refrigerators, automobiles and radios included in the index was improving but there was no way to measure the extent of the improvement of the models priced from year to year, and to that extent the index has had an upward bias.

(d) Mr. Teper raises the question as to whether, with wartime changes, the cost of living index measures currently changes in the cost of living or in the plane of living as well. When new automobiles and other durable goods were dropped from the index, beginning in January 1942, no substitutions were made except for increases in local transportation and automobile repairs, and the durable goods were "linked out." The level of the index did not decline because of the deletion of these articles. Mr. Teper suggests that the Bureau should preserve the level of consumption priced in pre-war days and allow equivalent commodities and services to replace those not obtainable. This is a difficult question, for the goods are not available; there is nothing equivalent to replace them. Thus, currently the Bureau has sought to account for price changes in goods actually available on the market.

We have frequently urged that the cost of living index should be accompanied by figures on changes in family expenditures during the war period. Funds have not been made available for such studies, and the Bureau has thus been unable to furnish any measure of changes in the plane of living since March 1942.

(e) Finally, the Bureau agrees that Mr. Teper's point is well taken that it is desirable to provide both the general public and technicians with more information about the way in which the cost of living index is constructed and how it can be used. Mr. Teper has been assured that, with added funds just made available for more extended work on prices and the cost of living, the Bureau will make every effort adequately to publicize the cost of living index.

ANALYSIS OF VARIANCE FOR PERCENTAGES BASED ON UNEQUAL NUMBERS*

BY W. G. COCHRAN
Iowa State College

SEVERAL WRITERS have discussed the use of the analysis of variance when the data are expressed as fractions or percentages. In the case where all percentages are based on the same total count (i.e., the same demoninator), the techniques which have been developed appear to be adequate for most practical purposes. According to the nature of the variations present in the data, these techniques take the form of (i) a straightforward analysis of variance of the percentages, (ii) an analysis of variance of *angles* into which the percentages have been transformed, [Bliss (1938), Cochran (1938), Clark and Leonard (1939)] or (iii) the angular transformation with some further refinements [Bartlett (1936) (1937), Cochran (1940)].

The case in which the percentages are based on different total numbers is more troublesome. As will appear later, there is a greater variety of possible methods of analysis, so that much time may be consumed in trying to decide what is a suitable method. Further, with certain types of data the efficient methods are rather tedious. The object of this discussion is to suggest approximate preliminary tests which are helpful in the choice of a method of analysis that is reasonably efficient and not unnecessarily laborious.

NATURE OF THE VARIATION IN PERCENTAGE DATA

For convenience in this discussion fractions will be used instead of percentages. Let n_i be the total count on which an observed fraction or percentage is based, and let $f_i = a_i/n_i$ be the observed *fraction*. If p_i is the true fraction for this observation, the variance of f_i about p_i is given by the usual expression for the binomial distribution: p_iq_i/n_i, where $q_i = (1-p_i)$.

In addition, the true fraction p_i will frequently be found to vary from observation to observation. Where present, such variation may contribute to the experimental error on which z, F, or t-tests are made. Thus for the total variance of an observed fraction we may write

$$V(f_i) = \frac{p_iq_i}{n_i} + \sigma_i^2 \tag{1}$$

---

* Journal paper No. J-1115 of the Iowa Agricultural Experiment Station, Ames, Iowa. Project No. 514.

where the first term represents the *binomial* variation and the second term the *extraneous* variation.

It is not clear what assumptions may best be made about σ_i^2. Probably no single set of assumptions is appropriate for all types of data. Since, however, σ_i^2 represents the variance of the *true* fraction, σ_i^2 will not in general depend on n_i. It may depend on p_i, for more variation would be expected when the fractions are near ½ than when they approach 0 or 1.

Subject to the assumption, implicit in all uses of the analysis of variance with fractions, that the fractions are approximately normally distributed, equation (1) suggests that the observations should be *weighted*, where the true weight W_i is

$$W_i = \frac{1}{\dfrac{p_iq_i}{n_i} + \sigma_i^2} = \frac{n_i}{p_iq_i + n_i\sigma_i^2}. \tag{2}$$

When the extraneous variation is small, the weight assigned to any fraction is n_i/p_iq_i; i.e., it increases directly as the total n_i from which the fraction is derived. At the other extreme, when the values of σ_i^2 are large relative to p_iq_i, the weight of an observation is $1/\sigma_i^2$, i.e., independent of n_i.

The practical procedure which follows from this discussion is to estimate first the relative amounts of binomial and extraneous variation in the data and the nature of the extraneous variation. From this information an appropriate set of weights can be constructed for the subsequent analysis. But unless the data are extensive or unless *a priori* knowledge is available, it may be expected that the information for the calculation of weights will not be very accurate. Moreover the computation of weights is time-consuming. For these reasons it is worthwhile to consider the adequacy of some simple methods of weighting.

THE EFFICIENCIES OF BINOMIAL AND EQUAL WEIGHTING

Two methods to be investigated are (i) weighting proportional to the n_i (ii) weighting independent of the n_i. As we have seen, these methods are efficient when the ratio of the binomial to the total variation is 1 and 0 respectively.

For simplicity we will consider the estimation of the mean of a group of s observations, such as the mean for one treatment over a number of replications. If weights w_i are used where the true weights are W_i, the variance of the weighted mean is known to be

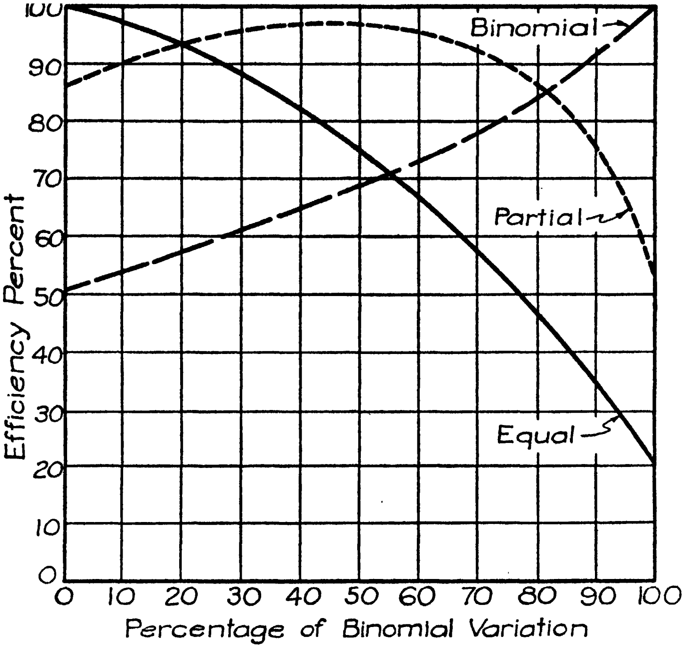
$$\sum_{i=1}^{s}\left(\frac{w_i^2}{W_i}\right) \Big/ \left[\sum_{i=1}^{s}(w_i)\right]^2. \tag{3}$$

Since the variance of a correctly-weighted mean is $1/\Sigma^{s}_{i=1}(W_i)$, the efficiency of the actual mean is

$$\left[\sum_{i=1}^{s}(w_i)\right]^2 \Big/ \left[\sum_{i=1}^{s} W_i\right]\left[\sum_{i=1}^{s}\left(\frac{w_i^2}{W_i}\right)\right]. \tag{4}$$

In order to obtain data which can be visualized, a number of sets of values of n_i were chosen, in some of which the n_i were approximately equal while in others they varied greatly. It was assumed that (p_iq_i) and σ_i^2 were constant for all observations. This assumption is not considered unduly restrictive since the mean of a *single* treatment is being considered. The assumption might fail if, for instance, a treatment in a randomized blocks experiment gave 95 per cent in some blocks and 40 per cent in others.

EFFICIENCIES OF BINOMIAL, EQUAL AND PARTIAL WEIGHTS



With these assumptions we may take $w_i = n_i$ (binomial weighting) in method (i) and $w_i = 1$ (equal weighting) in method (ii). For each of the selected sets of n_i, the efficiencies of the two weightings were calculated by formula (4). These efficiencies naturally depend on the ratio of the binomial to the total variation. Since the average binomial variance is

$pq/\bar{n}_h$, where $\bar{n}_h$ is the *harmonic* mean of the n_i, the ratio of the binomial to the total variance was taken as

$$u = \frac{pq}{\bar{n}_h} \Big/ \left(\frac{pq}{\bar{n}_h} + \sigma^2\right). \tag{5}$$

Each efficiency was computed for a series of values of u, ranging from 0 to 1.

As would be expected, the efficiencies of both methods were always close to unity when the n_i varied little. In this case either weighting is satisfactory whatever the relative amount of extraneous variation, equal weighting being preferable on account of its simplicity. For a case of extreme variation in the n_i, the efficiencies are shown graphically in the Chart. The third curve (partial weighting) will be explained later. The values of n_i were 1, 2, 6, 8, 10, 20, 30, 40, 50 and 80, exhibiting a range which is seldom equalled or exceeded in experimental data.

In the Chart binomial weighting is seen to be more accurate when the binomial variation in the data exceeds about 55 per cent and is less accurate otherwise. Binomial weighting at its worst ($u=0$) is superior to equal weighting at its worst ($u=1$). Since this result is always true, binomial weighting would appear to be a better method for routine use, there being less chance of getting a very inefficient mean than with equal weights. Neither method is however satisfactory in the mid-range when the binomial variation lies between 30 and 80 per cent of the total variation.

It should be remembered that the example represents a very extreme case. For examples which are perhaps more typical, in which n_i varied over a much less extreme range than used for the illustration shown in the Chart, the efficiency of binomial weighting appears to exceed 80 per cent throughout the range, while that of equal weighting may fall to about 60 per cent. Nevertheless there is need for a better method when about half the variation is binomial and the n_i are far from constant.

PARTIAL WEIGHTING

In the numerical examples which were worked, examinations of the correct weights as given by formula (2) revealed that when about half the variance is extraneous, the weights tend to change slowly for the higher values of n_i. This feature is illustrated in Table I, where the correct weights are compared with the values of n_i in the example of the Chart in the case in which half the variation is extraneous.

All W_i were multiplied by a common factor to make the true and binomial weight the same (unity) for the first observation. It will be

TABLE I

BINOMIAL WEIGHTS (n_i) AND CORRECT WEIGHTS WHEN $u=0.5$

n_i	1	2	6	8	10	20	30	40	50	80
W_i	1	1.71	3.26	3.67	3.98	4.76	5.10	5.29	5.41	5.59

noted that the increase in W_i for a fixed increase in n_i becomes steadily and rapidly smaller as n_i increases. Thus when n_i increases from 1 to 2, W_i increases from 1 to 1.71; yet when n_i increases from 50 to 80, W_i increases only from 5.41 to 5.59. In fact, the extraneous variation imposes a ceiling on the weight assigned to any observation, no matter how high n_i is.

In these circumstances a simple method of approximating the correct weights is to assign a fixed upper limit to the weights. Below that limit the n_i are chosen as weights. This device, which may be called *partial* weighting, was previously suggested [Cochran (1937), Yates and Cochran (1938)] for an analogous problem. At first a weighting was tried in which the upper limit was chosen as the average of the n_i (24.7). With this rule the values of w_i are 1, 2, 6, 8, 10, 20, 25, 25, 25, 25 respectively. The examples indicated, however, that this weighting reaches its highest efficiency when there is between 80 and 90 per cent binomial variation. The best simple convention for u between 30 and 80 per cent is to assign a constant weight to the upper two-thirds of the n_i. For the example cited these weights would be 1, 2, 6, 8, 8, 8, 8, 8, 8, 8. In computation this procedure has the additional advantage of giving fewer unequal weights to deal with.

The efficiency of this type of partial weighting is shown by the dotted curve in the Chart. When u lies between 20 and 80 per cent, partial weighting is superior to both binomial and equal weighting. Its efficiency is above 90 per cent for any value of u between 10 and 70 per cent.

Partial weighting has an element of arbitrariness in the choice of the upper limit. The choice appears to make relatively little difference if somewhere between one-half and two-thirds of the observations are given the same weight. The latter rule is preferable for general use.

PRELIMINARY INVESTIGATION OF DATA TO BE ANALYZED

We now discuss certain preliminary calculations which assist in deciding how a batch of data is to be analyzed. The first question to be considered is: are the variations in the totals n_i sufficiently small so that they can be ignored? Sometimes the answer, one way or the other, is obvious on inspection; but frequently with moderate variation in the totals it is doubtful whether the extra work involved in a weighted analysis is justified.

If variations in the n_i are ignored when they should be taken into account as in formula (2), some loss of efficiency results (and in addition some disturbance of the significance levels for z, F, or t-tests). From the preceding discussion it is clear that this loss cannot properly be estimated without knowledge of the relative amounts of binomial and extraneous variation in the data.

An upper limit to the loss can however be assigned from the values of the n_i. For it may be shown from formula (2) that the loss is greatest when all the variation is binomial. (The Chart provides an illustration of this result.) Consequently an upper limit to the loss is obtained by calculating the efficiency obtained from equal weights on the assumption that all the variation is binomial.

The purpose of the analysis of variance is presumed to be to estimate and test certain group or treatment averages. As might be expected, the potential loss of efficiency from equal weights depends on the type of classification present in the data (e.g., single grouping, two-fold classification or multiple classification.) Only the simplest and most common types—the single grouping and the two-fold classification—will be considered here.

For the same data the loss differs also from one type of treatment comparison to another. An *average* figure may be obtained by supposing that the true treatment means are distributed about the general mean with variance σ_t^2 and that this variance is the quantity to be estimated.

Single grouping. With r treatments and s observed fractions for each treatment, the totals n_{ij} on which the fractions are based may be set out as in Table II.

TABLE II

	Treatments				
	1	2	$\cdots$	r	
	n_{11}	n_{21}	$\cdots$	n_{r1}	
	n_{12}	n_{22}	$\cdots$	n_{r2}	
	$\cdot$	$\cdot$	$\cdots$	$\cdot$	
	n_{1s}	n_{2s}	$\cdots$	n_{rs}	
Totals	$N_{1\cdot}$	$N_{2\cdot}$	$\cdots$	$N_{r\cdot}$	N

When reduced to a comparable basis, the expected values of the treatments mean square for the two types of analysis work as follows:

$$\textit{Binomial weights}: pq + \frac{1}{(r-1)}\left[N - \frac{N_{1\cdot}^2 + \cdots + N_{r\cdot}^2}{N}\right]\sigma_t^2. \quad (6)$$

The coefficient of σ_t^2 is always slightly less than the mean of the treatment totals $N_{i.}$ (unless these totals are identical). If the number of treatments exceeds 10, or if the $N_{i.}$ vary little, the mean of the $N_{i.}$ may be used in (6) as a sufficiently good approximation.

Equal weights:

$$pq + s\bar{n}_h\sigma_t^2 \tag{7}$$

where $\bar{n}_h$ is the harmonic mean of the n_{ij}, given by the equation

$$\frac{1}{\bar{n}_h} = \frac{1}{rs}\left(\frac{1}{n_{11}} + \cdots + \frac{1}{n_{rs}}\right). \tag{8}$$

For the estimation of σ_t^2 the lower limit to the efficiency of equal weighting is therefore

$$s(r-1)\bar{n}_h \Big/ \left[N - \frac{N_{1.}^2 + \cdots + N_{r.}^2}{N}\right]. \tag{9}$$

If this ratio exceeds 0.9 it may be concluded that variations in the n_{ij} can be neglected unless the data require the best possible analysis.

Equal weights within treatments. It sometimes happens that the totals n_{ij} are substantially constant within each treatment, but change considerably from one treatment to another. In this case, although a completely equal weighting is possibly unsatisfactory, an analysis in which each observation is weighted by the average n_{ij} for that treatment may be highly efficient. The lower limit to the efficiency of this type of weighting relative to binomial weighting may be shown to be

$$(r-1)s^2/\lambda \tag{10}$$

where λ is the rather complex expression

$$\lambda = \left[\left(N_{1.} - \frac{N_{1.}^2}{N}\right)\left(\frac{1}{n_{11}} + \cdots + \frac{1}{n_{1s}}\right) + \cdots + \left(N_{r.} - \frac{N_{r.}^2}{N}\right)\left(\frac{1}{n_{r1}} + \cdots + \frac{1}{n_{rs}}\right)\right]. \tag{11}$$

A close approximation to (10) which is more quickly calculated is

$$rs^2/\mu \tag{12}$$

where

$$\mu = \left[N_{1.}\left(\frac{1}{n_{11}} + \cdots + \frac{1}{n_{1s}}\right) + \cdots + N_{r.}\left(\frac{1}{n_{r1}} + \cdots + \frac{1}{n_{rs}}\right)\right]. \tag{13}$$

TABLE III

NUMBER OF SERVICES (n), NUMBER OF CONCEPTIONS (a) AND PERCENTAGE OF CONCEPTIONS TO SERVICES IN ARTIFICIAL INSEMINATION TESTS

Sample	Bull																	
	1			2			3			4			5			6		
	n	a	%	n	a	%	n	a	%	n	a	%	n	a	%	n	a	%
1	13	6	46	10	7	70	23	12	52	3	1	33	12	5	42	17	6	35
2	13	4	31	8	3	38	16	7	44	15	7	47	11	7	64	22	15	68
3	19	7	37	9	4	44	14	8	57	9	3	33	16	8	50	39	23	59
4	6	3	50	8	5	62	15	6	40	14	3	21	16	11	69	16	6	38
5	13	8	62	1	1	100	6	5	83	10	7	70	13	10	77	21	12	57
6	9	3	33	2	0	0	15	10	67	5	4	80	16	13	81	7	6	86
7	2	0	0	5	1	20	3	1	33	4	3	75	15	13	87	17	13	76
8	9	5	56	1	0	0	6	3	50	8	4	50	8	5	62	14	8	57
9	10	3	30	8	5	62	14	9	64	24	11	46	8	6	75	34	10	29
10	7	3	43	17	10	59	23	16	70	21	3	14	2	2	100	30	18	60
Total	101	42	42	69	36	52	135	77	57	113	46	41	117	80	68	217	117	54
Σ(rec.)	1.4151			3.3449			1.0924			1.3468			1.2553			0.5758		

Grand totals: (n), 752; (a), 398; (rec.), 9.0303

NUMERICAL EXAMPLE

The practical use of these formulae may be illustrated by a numerical example.[1] In this investigation a number of semen samples were taken from each of a number of bulls. With each sample one or more cows were inseminated artificially. The data in Table III show the number of inseminations (n_{ij}) and the number of conceptions (a_{ij}) for every sample. The fraction a_{ij}/n_{ij} measures the success of the sample in artificial insemination and is the variable to be studied. As a preliminary step in the investigation it is desired to test the differences among bulls. The original data contain the results for 21 samples for each of 17 bulls. For illustration 6 bulls were chosen, and for each bull 10 of the 21 samples were selected at random.

The numbers of services fluctuate widely both within and among bulls. If nearly all the variation is binomial, neither equal weighting nor equal weighting within treatments appears satisfactory at first sight, though the latter would be expected to be somewhat more efficient than the former.

The first step in applying the preliminary tests is to record the sum of the reciprocals of the n_{ij} for each treatment (i.e. bull), as shown below the treatment totals (1.4151 etc.). With the aid of a table of reciprocals these figures can be summed directly on the calculating machine. Their total over all treatments is 9.0303.

---

[1] I am indebted to Dr. G. W. Salisbury, Department of Animal Husbandry, Cornell University, for permission to use these data.

Three auxiliary quantities are now calculated as follows:

(i) $$N - (N_{1.}^2 + \cdots + N_{r.}^2)/N$$
$$= 752 - [(101)^2 + \cdots + (217)^2]/752 = 610.07$$

(ii) $$\bar{n}_h = 60/(9.0303) \qquad = 6.644$$

where $60 = rs$ is the total number of samples tested.

(iii) $\mu = 101 \times 1.4151 + 69 \times 3.3449 + \cdots + 217 \times 0.5758 = 945.2.$

The last quantity is required for the approximate test in formula (12) of equal weights within groups and need not be calculated if the treatment totals $N_{i.}$ are all approximately the same.

From formula (9), the efficiency for equal weights is

$$10 \times 5 \times 6.644/610.07 = .544.$$

From formula (12), the efficiency for equal weights within groups is approximately

$$6 \times 100/945.2 = .635.$$

The exact value, by the more laborious formula (10) is found to be 0.628.

Thus if the variation within bulls is entirely binomial, the analyses by equal weights and by equal weights within treatments are equivalent to discarding nearly 50 and 40 per cent of the information respectively, a procedure which is unjustified for anything but a crude examination. This result does not however dispose of the question of weighting. Since the samples from the same bull were tested on different cows at different times and in different places, there is no *a priori* reason to suppose that the probability of securing a conception is constant for all samples from a single bull. The estimation of the percentage of extraneous variation, which requires an analysis of variance of the actual percentages of conceptions, constitutes the next step.

ESTIMATION OF THE RELATIVE AMOUNTS OF BINOMIAL AND EXTRANEOUS VARIATION

Either equal weights or binomial weights may be used. If on inspection most of the variation appears to be extraneous the former are advisable, while if most of the variation appears to be binomial the latter are preferable unless the weighted analysis is considered too complicated. The reasons for these choices are: (i) theoretically we expect a better estimate when the weighting used is closer to the correct weighting and (ii) if the preliminary guess turns out to be approximately correct, the preliminary analysis may be satisfactory for tests of significance. In this case no further analysis will be needed.

When extraneous variation is present, the error mean square is composed of binomial variation, extraneous variation and random fluctuations. For a simple group comparison, the expected value of the error mean square of the *unweighted* fractions may be taken as

$$E(s^2) = \frac{pq}{\bar{n}_h} + \sigma_e^2 \tag{14}$$

where p is the average of the true fractions and σ_e^2 is the average extraneous variance. Actually this result is not quite correct, since the expectation depends on the unknown values of the true fractions p_{ij} for each observation. In equation (14) the average fraction p has been substituted for these unknowns. The result should be sufficiently accurate for its purpose.

Thus if s^2 is the error mean square, the ratio of the binomial to the total variance is estimated as

$$\bar{f}(1 - \bar{f})/s^2\bar{n}_h, \tag{15}$$

where $\bar{f}$ is the mean observed fraction.

If a *weighted* analysis is used, with weights n_{ij}, the error mean square s_w^2 is an estimate of

$$pq + \frac{\sigma_e^2}{r(s-1)}\left[N - \frac{(n_{11}^2 + \cdots + n_{1s}^2)}{N_{1\cdot}} \cdots \frac{(n_{r1}^2 + \cdots + n_{rs}^2)}{N_{r\cdot}}\right]. \tag{16}$$

The coefficient of σ_e^2 is slightly smaller than the arithmetic mean $\bar{n}$ of the n_{ij}. Unless the number of observations or replications per treatment is less than 4, it is usually sufficiently accurate to use $\bar{n}$ in (16), writing

$$E(s_w^2) = pq + \bar{n}\sigma_e^2. \tag{17}$$

Since the average binomial variance has been defined as $pq/\bar{n}_h$, the estimates of the two components of variance are:

$$\textit{Binomial}: \frac{\bar{f}(1 - \bar{f})}{\bar{n}_h}; \qquad \textit{Extraneous}: \frac{s_w^2 - \bar{f}(1 - \bar{f})}{\bar{n}}. \tag{18}$$

As a rough working rule it is suggested that if less than 30 per cent of the variation is binomial, equal weights may be employed for the analysis on which tests of significance are based. Binomial weighting is satisfactory if more than 80 per cent of the variation is binomial. In the intervening range partial weights may be used; or if the data are sufficiently extensive and important to warrant the additional work, the actual weights W_i in formula (2) may be estimated from the computations which have been carried out.

It should however be remembered that unless the number of degrees of freedom for error is large, the estimate of the percentage of binomial to total variation is itself subject to a large sampling fluctuation. Even if p were known exactly, the estimated percentage of binomial variation would at best be distributed as eB/χ^2, where B is the true percentage of binomial variation, e is the number of error degrees of freedom, and χ^2 has e degrees of freedom. For example, if a batch of data shows 50 per cent of binomial variation and the error has 20 degrees of freedom, the 95 per cent lower and upper limits to the true percentage are at least as far apart as 27 per cent and 78 per cent. With 30 degrees of freedom for error, this range narrows to 30–73 per cent. It may be concluded that with the moderate numbers of degrees of freedom which are common in experimental work, attempts to estimate the true weights W_i are scarcely justified, though the technique probably does permit the discrimination suggested between equal weights, binomial weights and partial weights.

APPLICATION TO THE NUMERICAL EXAMPLE

In order to form an initial judgment about the extent of extraneous variation, the percentages of conceptions in Table III were inspected. Although the percentages vary widely for the same bull when the number of inseminations is small, those percentages which are based on larger numbers of inseminations seem fairly stable. It was concluded as a first guess that the extraneous variation is small relative to the binomial variation. A *weighted* analysis was therefore made.

This analysis can be carried out on the fractions by the familiar "(total)^2/(number)" rule. For the sums of squares, we have

$$\textit{Total:}\qquad \frac{6^2}{13}+\frac{4^2}{13}+\cdots+\frac{18^2}{30}-\frac{398^2}{752}=24.058$$

$$\textit{Among bulls:}\qquad \frac{42^2}{101}+\cdots+\frac{117^2}{217}-\frac{398^2}{752}=\ 6.033.$$

The error (i.e., within bulls) mean square is 0.334 with 54 degrees of freedom.

The additional figures required are:

$$\bar{n} = 752/60 = 12.53 : \bar{n}_h = 6.644 : \bar{f} = 398/752 = .5293$$

$$\textit{Binomial variance} = \bar{f}(1-\bar{f})/\bar{n}_h = .0375$$

$$\textit{Extraneous variance} = [s_w^2 - \bar{f}(1-\bar{f})]/\bar{n} = .0068.$$

The ratio of binomial to total variance is estimated as 375/443, or about 85 per cent. From this result binomial weighting appears adequate.

To illustrate the alternative approach, the following is the *unweighted* analysis of variance of the percentages.

	d.f.	*Sums of squares*	*Mean squares*
Among bulls	5	6,274	1255
Within bulls (error)	54	24,144	447

The unweighted mean of the percentages is 52.42. From formula (15), this gives (52.42)(47.58)/(447)(6.644), or 84 per cent of binomial variation. The agreement between the two estimates is closer than will normally be found.

Further evidence on the superiority of binomial over equal weighting is provided by the F-ratios, which are 3.61 for the former and 2.80 for the latter.

THE ANGULAR TRANSFORMATION

Thus far nothing has been said about the use of the angular transformation, in which the analysis of variance is performed on the transformed variate $y = \sin^{-1}(\sqrt{f})$. For it seems best to settle the question of weighting before considering the advisability of a transformation.

The angular transformation was devised for the case in which all the variation is binomial with constant n. The effect of the transformation is to change the weights $n/(p_iq_i)$, which apply to f, into a set of approximately *constant* weights $n/821$ applicable to the angles y (expressed in degrees).

When the n_i vary and extraneous variation is present, the effect of the transformation is more complicated. Suppose however that the extraneous variance σ_i^2 is of the form (λp_iq_i). The correct weights for f and for y are then:

$$W_i(f_i) = \left(\frac{n_i}{1+\lambda n_i}\right)\left(\frac{1}{p_iq_i}\right) : W_i(y_i) = \left(\frac{n_i}{1+\lambda n_i}\right)\left(\frac{1}{821}\right). \quad (19)$$

These results illustrate the fact that the transformation equalizes approximately any system of variances which are proportional to

$(p_i q_i)$, but does not alter the extent to which the weights vary with the totals n_i. It is for this reason that we decide first how the weights shall vary with n_i. If an angular transformation is subsequently made, the weighting adopted should be used with the angles.

The relation $\sigma_i^2 = \lambda p_i q_i$ imposes a restrictive assumption on the nature of the extraneous variation. The assumption is perhaps not unreasonable for data where the variance is greatest at 50 per cent and smaller towards both ends of the range.

It seems therefore that the usual rules for deciding whether to use angles may be applied after the weights (equal, binomial or partial) have been chosen. If the percentages vary widely, especially as between different treatments, and if the error variance appears to be greatest around 50 per cent, the transformation may provide a more homogeneous estimate of error. In the numerical example a transformation was not considered necessary in view of the fact that the average percentage of conceptions for the different bulls ranged only from 41 per cent to 68 per cent.

TWO-WAY CLASSIFICATIONS

The principles outlined above apply also to two-way classifications (as in a randomized block experiment), though the detailed application is more complicated. The chief practical change is an increase in the attractiveness of equal weighting. With a given amount of variation in the n_{ij}, the efficiency of equal weighting is usually somewhat greater than in the case of a simple grouping. In addition, it has been shown (Yates, 1934) that a weighted analysis cannot be carried out by the elementary methods of the analysis of variance, because the sums of squares for the rows and columns of the classification are in general non-orthogonal. In order to test either rows or columns against their interaction, a set of multiple regression equations must be solved. For this calculation the procedure described for a continuous variate by Yates (1934) may be followed, provided that the weights w_{ij} are substituted for Yates's totals n_{ij}. Except in special cases, there will be no "within-classes" sum of squares.

Some modifications are required in the formulae used in the preliminary tests. According to the nature of the data, we may wish to consider one of three simple methods of weighting, each of which avoids the labor of a least squares solution: (1) equal weights; (2) equal weights within treatments; (3) equal weights within *rows* or *replications*. On the assumption that all the variation is binomial, the minimum values for the efficiencies of these methods are as follows:

Equal weights: $s(r-1)\bar{n}_h/\nu$ (20)

Equal weights within treatments:

$$rs^2[N-(N_{1\cdot}^2+\cdots+N_{r\cdot}^2)/N]/\mu\nu \tag{21}$$

Equal weights within replicates: $r(r-1)N^2/\theta\nu$ (22)

where

$$\nu = N - \frac{(n_{11}^2+\cdots+n_{r1}^2)}{N_{\cdot 1}} - \cdots - \frac{(n_{1s}^2+\cdots+n_{rs}^2)}{N_{\cdot s}} \tag{23}$$

$$\begin{aligned}\theta = N_{\cdot 1}^2\left(\frac{1}{n_{11}}+\cdots+\frac{1}{n_{r1}}\right)+\cdots \\ +N_{\cdot s}^2\left(\frac{1}{n_{1s}}+\cdots+\frac{1}{n_{rs}}\right).\end{aligned} \tag{24}$$

The quantities ν and θ involve respectively sums of the squares of n_{ij} and of the reciprocals of n_{ij} taken along the *rows*. Formula (21) is an approximation, μ having the value defined in formula (13).

The use of any of these formulae requires a certain amount of preliminary calculation. Where data of this type are frequently encountered, the computer soon becomes experienced in estimating the efficiencies simply by inspection of the values of the n_{ij}, so that in course of time the calculations need seldom be made.

Should none of these methods be considered satisfactory, the relative amounts of binomial and extraneous variation must be estimated. Unless it appears fairly certain that the greater part of the variation is binomial, estimation from an equally weighted analysis of variance is preferable on account of the saving in time. With equal weights, formula (14) remains valid for the expected value of the error mean square (in this case the treatments times replicates interaction). When binomial weights are used and the least squares analysis is carried out, the expected value of the error mean square is cumbersome. The approximate formula (17) appears to be sufficiently accurate for practical use; it underestimates slightly the relative amount of binomial variation.

SUMMARY OF SUGGESTED PROCEDURE

1. Consider whether one of the simplest methods of analysis (equal weights, equal weights within treatments, equal weights within replicates) can be used without further investigation. If in doubt apply the appropriate test.

2. If none of the methods in (1) appears satisfactory, estimate the relative amounts of binomial and extraneous variation from an analysis of variance of the fractions or percentages.

3. As a rough discriminatory rule, adopt binomial weights for the subsequent analysis if more than 80 per cent of the variation is binomial, and equal weights if less than 30 per cent of the variation is binomial. In the intervening range use partial weighting unless the data are extensive and important enough to justify the estimation of the exact weights for every observation.

With a two-way classification it may be advisable under pressure of time to weight equally when as much as 50 per cent of the variation is binomial. In the examples which I have investigated, the loss of efficiency in this case was less than half the loss when all the variation is binomial.

4. When a method of weighting has been adopted, consider whether an angular transformation is advisable.

This procedure can be compressed as experience accumulates.

A discussion of the effects of erroneous weights on the validity of z, F or t-tests is beyond the scope of this paper. It appears that the more closely the weights used approach the true weights, the smaller is the disturbance in levels of significance. Consequently a procedure designed to choose weights that are approximately correct should also lead to an analysis of variance for which the tabulated significance levels are approximately valid.

REFERENCES

Bartlett (1936). *Journal of the Royal Statistical Society*, Supplement, 3, 68.
Bartlett (1937). *Journal of the Royal Statistical Society*, Supplement, 4, 168.
Bliss (1938). *Ohio Journal of Science*, 38, 9.
Clark and Leonard (1939). *Journal of the American Society of Agronomy*, 31, 55.
Cochran (1937). *Journal of the Royal Statistical Society*, Supplement, 4, 102.
Cochran (1938). *Empire Journal of Experimental Agriculture*, 6, 157.
Cochran (1940). *Annals of Mathematical Statistics*, 11, 335.
Yates (1934). This JOURNAL, 29, 60.
Yates and Cochran (1938). *Journal of Agricultural Science*, 28, 556.

"RETRO" CHARTS

BY KARL KARSTEN AND EDITH BROOKS

IT IS THE PURPOSE of this paper to introduce a family of useful chart-forms, discussing the necessary and sufficient conditions for their use, the main features of these charts, the different variations which are possible, and the use of these charts in collections where standardization is desirable.[1]

A census[2] of chart-literature indicates that the majority of widely-circulated charts show curves of time-series, and that in the greater part of these the reader is expected to be more interested in recent events than in the remote past. In this large group of charts the chart-maker can be said to play the role of historian.

And here the folkways of chart-making sometimes confine us in methodological rigidities which are absurd. Wouldn't it be silly if the historian were forced to cover a uniform number of years per page throughout his book? And if we are making a historical chart, may there not be occasions when it is just as silly for us to use equal distances for equal units of time? When such rigidity causes an error in the placing of emphasis, it certainly makes a poor chart. If the error causes deception, it makes a misleading chart. Both results are explainable, but not necessarily excusable, on grounds of convenience or of custom.

Custom has not so shackled the historian. For him it is accepted practice to devote much space to important periods of time and cover briefly what is unimportant. This treatment is realistic. It omits undesired details. It meets the reader's needs and interests.

The statistician often seeks escape from these rigidities which the historian has never known. He may try the device of breaking his time-series chart into two parts, one of which shows yearly data, for example, over a long period on a condensed scale, while the other part shows monthly figures, let us say, for only a few years on an enlarged scale. He gives both parts the same vertical scale, he joins them together at a common meeting-point of time without overlapping, and lo and behold, he has achieved great detail for the recent past and great coverage for the remote past.

---

[1] Previous contributions on this subject have been incomplete. See Karl Karsten, *Charts and Graphs* (Prentice-Hall, Inc., New York, 1925), pp. 474–5. The present writers are indebted to Mr. Thomas Blaisdell and Mr. Spurgeon Bell of the National Resources Planning Board for the opportunity to develop and explore the use of "retro" charts for collections of standardized charts. We are indebted to Col. Lewis Sanders for his insistence that we try out the simple perspective type of "retro" chart.

[2] One sampling of 274 charts from 26 publications (of trade conditions) showed 97 per cent were charts of historical time-series and only 3 per cent belonged to other types. Of the time-series charts, 63 per cent covered between 5 and 50 years and would be appropriate subjects for the present discussion;

The existence of such segmented double-gaited charts is eloquent evidence of the conflict between rigidities and requirements. But it is not a satisfactory solution of the conflict. The chart requires difficult mental acrobatics when reading the curve or curves from one part to the other.[3] It has other disadvantages. In view of the conflict, however, it is an ingenious compromise.

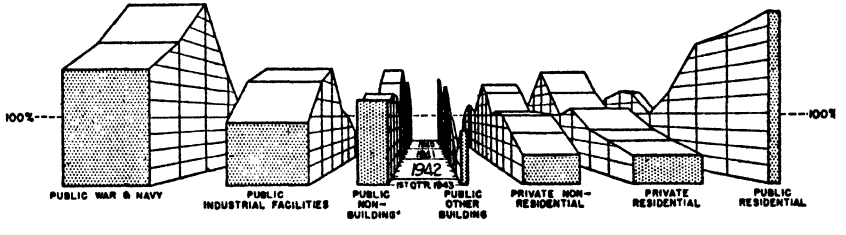
MAIN FEATURES

In many respects a better compromise is afforded by the family of "retro" or retrospective charts here proposed. These charts avoid the dreary expanse of the orthodox grid and the unbridged saltation of the segmented device. Like the latter, they differentiate between the recent and the remote past in detail and emphasis. Like the former, they preserve unbroken continuity between all adjacent portions of the chart. In short, they have the advantages of both and the disadvantages of neither.

CHART I

SEVEN "RETRO" CHARTS—PERSPECTIVE TYPE
CONVERTED INTO IMAGES OF SOLIDS BY USE OF A THIRD DIMENSION
(with Arithmetical Vertical Scale)

In this chart the indexes of the value of construction in the United States are shown for the past three years by seven curves each of which forms a separate retrospective chart.



Moreover, these charts introduce a graduated and continuous change in degree of emphasis and detail. Reading from right to left, we find the units of time continuously dwindling in size just as the memory of details fades, in life, with the passage of time. Strong argument can be offered that in this feature the "retro" chart provides a clearer and more truthful picture of past events seen in retrospect as a connected entirety, than is possible with the orthodox chart. It is not possible to design the "retro" chart without the implicit addition of a third dimension, which makes the past become dynamic, realistic, alive. (See Chart I.)

---

[3] It reminds one of the story by Jack London about a thief who evaded detection by shifting suddenly from the leisurely tread of a club member to the hurried pace of a waiter.

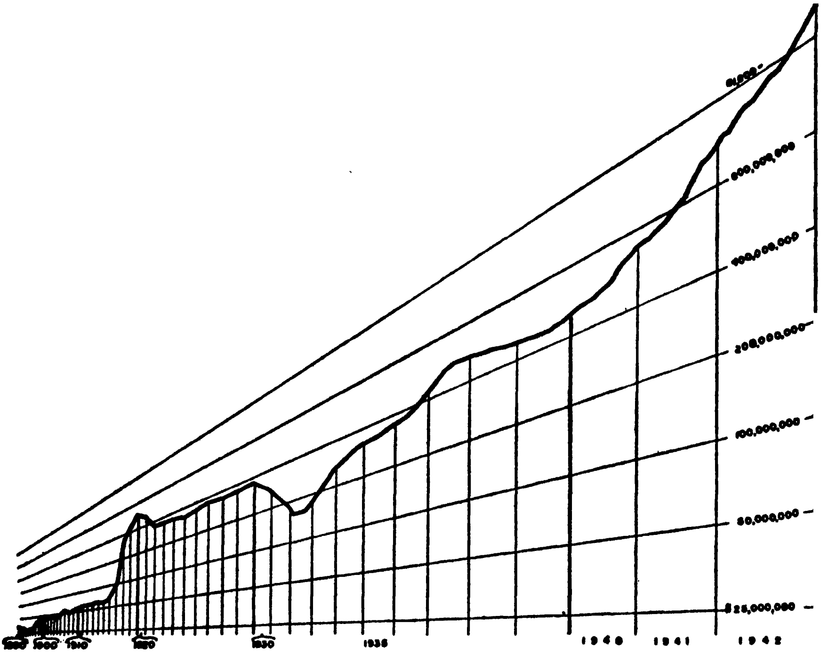
TRUE PERSPECTIVE

Conceptually the simplest "retro" chart is that which adheres faithfully to the rules of linear perspective and presents a perspective rendering in three dimensions. Imagine that the curve of the past has been drawn in the orthodox way on a very large scale, if you please, along the face of a high fence. Let us stand off from one end of this fence and view it from a position in front of the latter portion, just as Tom Sawyer must have seen the whitewashed fence as he stood off and admired it. At that moment the "retro" chart emerges, complete and inevitable. In it we view the past from one end only, as in life itself. (See Chart II.)

CHART II

"RETRO" CHART—PERSPECTIVE TYPE
(with Logarithmic Vertical Scale)

In this chart, showing average monthly Federal Government Receipts over the last 50 years, it is interesting to observe that without difficulty transition has been made from annual averages to quarterly moving averages in 1930, and from these to monthly moving averages in 1939.



Among the disadvantages of this chart is the fact that its grid is laborious to construct.[4] Moreover its vertical scale changes as one moves across the chart,[5] so that magnitudes must be plotted by inspection, or measured precisely along some abscissa where a ruler conveniently fits, and transferred to the proper abscissa by triangulation through the vanishing point.[6]

---

[4] To construct the perspective-"retro" scale, the following procedure may be used:

(1) Decide upon the total number, N, of time units to be shown, also the total space, S, in inches

As to the slopes of the curves, it is noteworthy that straight lines on orthodox charts are straight lines on these charts.[7] But the angle of slope is not the same for different straight lines which would be parallel on orthodox charts. The reader must obviously compare the slope of each rise or decline with its background grid (the slope of adjacent ordinate rulings) before reading significance into the slope of the curve. Measured from that background grid, the angle of slope has the same significance as in orthodox charts. (See Chart III.)

OTHER TRIANGULAR FORMS

Many variations of the triangular "retro" chart can be made, either on logical grounds or on grounds of convenience in drafting. These

---

to be covered by these time units on the "retro" scale, and the ratio, R, such that $1/N < R < 1.0$, between the space given to the most recent time unit and the total space, S. (This ratio, of course, determines the extent of variation of the "retro" scale from the orthodox scale.)

(2) Compute the constant; $Q = \frac{N - RN}{RN - 1.0}$ and the values of the following series; $a = 0$; $b = \frac{1}{Q+1}$; $c = \frac{2}{Q+2}$; $d = \frac{3}{Q+3} \cdots n = \frac{N}{Q+N}$ · (The constant, Q, corresponds to the usual distance from the eye to the picture in the perspective rendering; it serves to obtain the desired ratio, R, of variation from the orthodox scale.)

(3) Compute the vanishing point; $V = \frac{S}{N}(Q+N)$, the distance in inches along the "retro" scale from the most recent date to the vanishing point. This point must be noted on the drawing as all the radiating ordinates must be extended to pass through the radiate from this point, even though the point lies, of course, outside the finished chart.

(4) Measuring backward in inches along the "retro" scale from the most recent date to the vanishing point, plot the successive points of time at the values of aV, bV, cV, $dV, \ldots$ nV to form the desired "retro" scale.

(5) From the vanishing point, V, project the converging lines corresponding to the horizontal ordinates of the orthodox chart.

[5] The vertical scale has not been discussed in the present text because the "retro" chart whether rectangular or triangular can obviously have the same variety of vertical scales that orthodox charts have. In Chart I, the arithmetical vertical scale has been used while Charts II, III, and IV have logarithmic vertical scales. The regular spacing of ordinates for scale values in strict geometric progression employed on these examples is believed to be helpful in explaining logarithmic charts to the layman and is certainly a great advantage for the triangular "retro" charts since it avoids excessive crowding of converging ordinate lines. Credit for this refinement belongs to Mr. Spurgeon Bell.

[6] The term vanishing point is here used to denote the point where converging ordinates meet in all triangular "retro" charts. This point can, of course, be placed whenever one wishes along the X-axis in the case of some of the pseudo-perspective triangular "retro" charts discussed in the following paragraph, namely, in the logarithmic and the backward square-root scales. For the forward powers projection, the point of origin in the X-axis would serve as a logical converging point for the ordinate rulings.

[7] This is true whether the vertical scale be arithmetical or logarithmic, for it is obvious that the "retro" chart, which is but a perspective rendering of the orthodox chart, can have any vertical scale and still reproduce straight lines as straight lines.

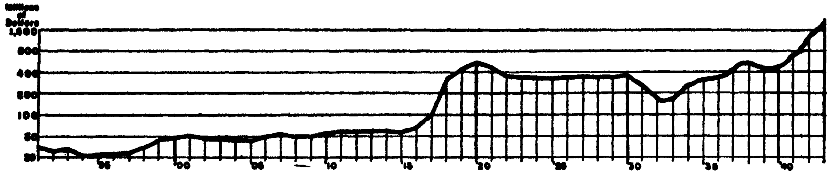
form a group which might be called "pseudo-perspective retro" charts. A backward logarithmic scale can be constructed very easily by counting the time-units backward from some arbitrary origin in the future and plotting the resulting values from a slide-rule.[8] A somewhat similar scale can be constructed by plotting the square-roots of these units.

CHART III

ORTHODOX CHART

(with Logarithmic Vertical Scale)

This chart of average monthly Federal Government Receipts is comparable to the preceding "retro" chart (Chart II) showing precisely the same figures plotted on the orthodox time scale. It is this chart which we can imagine portrayed in Chart II after considerable expansion followed by rotation through nearly 90 degrees.



Alternatively, the time-units can be counted forward from an assumed origin previous to the series, and plotted according to their squares or other powers, or according to their anti-logarithms. These and other devices have individual advantages and disadvantages, but in general serve to graduate the time-scale in what could be called a retrospective manner.[9]

RECTANGULAR FORMS

Between the foregoing triangular "retro" charts and the orthodox grid, there can be found a large group of intermediate compromises of rectangular shape. In these the feature of converging ordinates meeting at a vanishing point is abandoned and a return is made to the parallel horizontal ordinates of the orthodox grid. This is analogous to the isometric drawing prepared by architects and engineers whereby uniform scales are employed in flat pictures showing various surfaces of a three-dimensional solid. There results a sort of "mercator's projection" of the triangular "retro" chart. (See Chart IV.)

---

[8] If it be true that among the factors which control inherited characteristics, a parent is twice as important as a grandparent, four times as important as a great-grandparent, eight times as important as a great-great-grandparent, then a logical analogy is established for the use of the logarithmic scale, in which the years march backward at such diminishing pace that the space given to the two years before the last is the same as the next previous four years, and also to the next previous eight years, and back of that to the next previous sixteen years, and so on in a geometric progression.

[9] In most "retro" charts a certain point of time is treated as a point of origin or reference, and the time scale cannot be extended or extrapolated past this point any more than a curve on logarithmic rulings can pass from positive to negative values. In the perspective projection, and in the backward square-root and logarithmic projections, the chart can be extrapolated backward indefinitely, but not forward. In the forward antilog and powers projections the reverse limitation obtains, and the time scale can be extended or extrapolated forward to any extent but not backward. These differences will sometimes be important in determining the choice of type of "retro" chart to employ. It is also important to note that in the pseudo-perspective triangular "retro" charts, the power to reproduce as a straight line every straight line on orthodox charts has been lost.

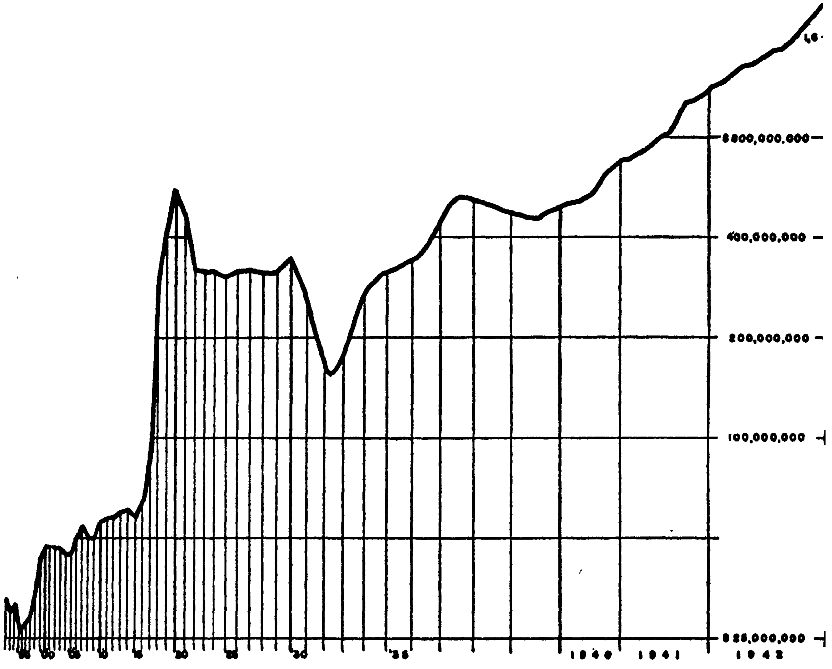
This group of hybrids is highly conventionalized, and is of value only when its conventional nature is fully and correctly understood. It attains uniformity of vertical scale across the chart, which is a great convenience in construction and in analytical reference. But it has sacrificed the quality of reproducing as a straight line any trend which would describe a straight line on an orthodox grid. Therefore the angles of slope cease to be comparable over different portions of the time-scale and become as misleading to the unwary as they are in the orthodox logarithmic chart for the arithmetically minded reader.[10]

CHART IV

"RETRO" CHART—ISOMETRIC TYPE

(with Logarithmic Vertical Scale)

Here is a third chart of the average Federal Government Receipts shown in Charts II, and III, but the horizontal rulings are parallel as in an orthodox chart. The retrospective scale used in the chart is a backward logarithmic projection such as can be easily plotted from a slide rule.



---

[10] For most persons these limitations of the "retro" chart will not be of much practical importance. Few readers of a chart concern themselves with the mathematical concepts of acceleration and decelera-

OTHER PATTERNS OF EMPHASIS

Lastly, there is a large group of charts analogous to "retro" charts which are not retrospective but achieve some other historical point of view. Thus charts could be designed to look into the future instead of into the past, and to see the distant future smaller than the near future. Better still, we may "panoram" from the past, through the present and into the future, in a single chart.[11] The graduations of such a chart might taper off at each end by a reverse function of the normal probabilities projection.[12] Or it may be desirable to expand in great detail two periods and compress an intermediate period as well as the exterior periods of time. Such a bi-modal form, high-lighting the first and second world war periods, might be analytically useful for war-time planning. These statistical refinements will vary with different students and studies. They will not, however, have universal historical value, nor will they achieve the most satisfactory simplicity, but for particular purposes they can provide a most powerful and illuminating chart-medium.

STANDARDIZATION BENEFITS

For use with other "retro" charts of the same type, especially when forming a large collection of charts with identical time-scales, to facilitate comparison, the "retro" chart becomes a great practical convenience. This is particularly true if—as is usually the case—the charts must extend backward to different dates because they show time-series which cover different lengths of time.[13]

Mathematically it would seem that by no other device could a wide variety of historical time-series be presented upon a single standard time-scale with satisfactory and proper individual detail for each one, when the series differ extremely in length.[14]

---

tion measured by successive differences and shown visually by the tangent of the curve. Also it should be remembered that the slopes of curves have only limited significance in any chart. This must be correctly understood by the reader or it may prove deceptive. Thus in the common arithmetically-scaled chart, the angle of slope is significant as to amount of change but deceptive as to rate of change, while on the ratio or semi-log chart the reverse is true, and the angle of slope is significant as to rate of change but may be most deceptive as to amount of change..

[11] This two-way historical chart or panorama can consist of two rectangular "retro" charts or of two triangular "retro" charts, in either case joined at the expanded portions of the scales. For very simple use, the double triangular form is preferable as it approximates the result obtained with a panoram camera.

[12] For an example of this projection, see Karsten, *op. cit.*, p. 457.

[13] In this respect its convenience for use on the horizontal scale reminds one of the convenience with which the logarithmic vertical scale served to plot series fluctuating through widely different magnitudes—a feature which appealed to many when log charts were first introduced.

[14] On the one hand, it will be found that a short time-series covering only the last five years, for

Such charts are easy to read because it is unnecessary for the reader to re-orient himself carefully for each new chart. Every time-period in the past is always located in the same position on every chart.[15]

Comparison of one chart with another is facilitated because any chart can be readily superimposed upon another. The reader can lay them before him with one above the other and compare them; he can hold them up to the light and look through the two of them; he can trace one curve upon the other mentally or physically. There is minimal trouble and no inaccuracy in these comparisons.

A disadvantage of "retro" charts is that they are not so suitable as the orthodox chart for the type of time-lag analysis which consists of shifting one chart physically to the right or left before comparing it with another. With these charts it becomes necessary to keep on shifting and readjusting the lagged chart as one compares it at different time-periods.[16] This can be performed with these charts by progressive readjustment.

One should not take too seriously the disadvantages and limitations which have been recited. They apply equally to the segmented double-gaited orthodox monstrosities which have been described, and which are not commonly quarrelled with.

Since the "retro" chart is a popular, not a research tool, its merit lies in uses for which it proves to be a more truthful and effective medium of

---

example, will be shown in great detail over about half of the chart page. In the examples given here (Charts II and IV) the last four years by reason of their wide spaces show monthly data, and before that eight years of quarterly data before the spacing becomes too narrow for any but annual figures. In the other extreme case, a time-series covering the last half century will be shown in its entirety upon the single chart page (by reason of the omission of detail in the picture of the earlier years) yet each of these two charts will be comparable with the other. It would be impossible on orthodox charts to show the 50-year chart on the same scale as the five-year chart unless either (1) the latter were compressed to a very small size and its monthly details omitted, or (2) the 50-year chart were made extremely long so as to extend over several or many pages. Either of these two solutions would obviously be so unsatisfactory that the effort to attain standardization from chart to chart would necessarily be abandoned. In the end, each chart would be presented on its own appropriate scale with no intercomparability from chart to chart.

[15] To facilitate cross comparisons, all identifying marks along the time scale should be identical in style and typography. It is, of course, not necessary for the scale marks on a short series to be carried back beyond where the curve begins, but the abscissae should be alike whenever they occur. It may be useful to use extra heavy rulings for special abscissae such as the divisions between decades. It may be especially helpful to adopt a method of shading a portion of the chart lying under the curve between certain significant abscissae, thus the area representing war years or the depression years might be invariably shaded in a characteristic and recognizable fashion. These landmarks would help the reader to orient himself on each chart.

[16] This is not difficult since one is only keeping the two backgrounds "in gear." But if a stationary picture is desired, it would be necessary to replot the curve with the appropriate time-lag or -lead. This limitation is likely to be over-rated. Actually this type of analysis is much less frequently practiced than we imagine. Few time-series have rational bases for anything beyond very short time-lags or -leads (see Karl Karsten, "The Theory of Quadrature in Economics," this JOURNAL, March 1924, pp. 14–29). It is the power to perform this analysis that we have been taught to prize.

information.[17] This is the case when its peculiar arrangements convey to the reader their message with just the degree of detail desired. To a busy executive, in such cases, a service of these charts, besides being more informative, will be a time-saver.

---

[17] Twenty years ago the present writer prepared blank pages for history note-books in which the years were printed down the edge of the page in a backward logarithmic progression. These forms were found to be useful for historical notes on the widest variety of topics. Whether one was filling in the chronology of Chinese art, of American history, of European architecture, or of geological cycles, the retrospective scale was found to make appropriate provisions for the various periods covered. On such a scale it is quite possible to plot the historical events back to the approximate time of the formation of the earth with conveniently indicated margins of probable error of such dates. In recent years there has appeared on the market a brightly colored "Histogram" of evolution for schoolroom purposes tracing in several yards of such backward logarithmic-scale retrospective charts, the history of different kinds of plant and animal life from the first evidences of such life down to the present time.

A MECHANICAL DETERMINATION OF CORRELATION COEFFICIENTS AND STANDARD DEVIATIONS

By John R. Platt
Physics Department, Northwestern University

In computing a bivariate correlation coefficient by the product moment method, it is common practice to plot a "scatter diagram" of the two variables on a graph as an aid in visualizing the problem. However, it is not generally appreciated that the correlation coefficient, r, can be obtained dynamically from such a scatter diagram, if the points on the graph are given physical weight and if the graph itself is a fairly rigid web such as a wire mesh, by finding the physical moments of inertia of the system about various axes.

This procedure is subject to experimental error, but with careful technique, r can be found with an accuracy of .02; and the standard deviations, σ_x and σ_y, of the two sets of values are obtained in the process with an accuracy of better than 1 per cent. After the "points" have been located in the "graph," the time it takes to determine the moments with a simple torsion pendulum is about 8 minutes, and the time to compute r, σ_x, and σ_y, is less than 5 minutes. The only arithmetic necessary is an initial determination of the means of each set of values, and the final substitution in simple formulae. Any number of pairs of values can be handled from 20 to 200 or more.

It is felt that this method can be used to advantage in statistical problems where punch-card and calculating machines are not available or where ease of computation outweighs in importance absolute accuracy of result. Probably, however, its especial value would be in teaching the principles of correlation to students; for in this method the physical interpretation of correlation—as a measure of scatter from a perfect prediction line—is never obscured by a maze of arithmetical details.

THEORY OF OPERATION

The standard deviation, σ_x, of a set of values, x_i, from a mean value $\bar{x}$, is given by

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}}$$

where N is the number of values in the set. The "product moment" correlation coefficient between two matched sets of values, x_i, y_i, is given by

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}.$$

The sums involved in these expressions are familiar in mechanics in the determination of the inertial ellipsoid for a set of unit mass points in the xy-plane connected by a rigid weightless web. If the mass points have the coordinates, x_i, y_i, then σ_x is just the "radius of gyration" of the system about the y-axis.

$$\sigma_x = \sqrt{\frac{Y}{NI}} \tag{1a}$$

and similarly

$$\sigma_y = \sqrt{\frac{X}{NI}} \tag{1b}$$

and r is given by

$$r \quad \frac{P_{xy}}{\sqrt{XY}}$$

I is the "moment of inertia" of unit mass at unit distance from the axis of rotation; X and Y are the "moments of inertia" of the system about axes respectively parallel to the x and y axes and passing through the center of gravity $(\bar{x}, \bar{y})$; P_{xy} is the so-called "product of inertia." The nomenclature of correlation was taken by Pearson directly from these dynamical terms in his first analysis of the theory of correlation;[1] but so far as the author knows, no previous use has been made of these dynamical analogies for actual computation of correlation coefficients. (Harsh and Stevens in their correlation machine utilize only the static analogy between the mean of a set of points and the center of gravity of a set of equal weights.[2]

P_{xy} is related to actual moments in several ways, the simplest connection being deduced from the formula

$$D = X \cos^2 \theta + Y \sin^2 \theta - 2P_{xy} \sin \theta \cos \theta$$

(where D is the moment of inertia about an axis in the xy-plane making an angle θ with the positive x-axis), by setting $\theta = 45°$. Then

---

[1] Karl Pearson, "Regression, Heredity, and Panmixia," *Philosophical Transactions of the Royal Society*, A, 187, 253 (1896).

[2] C. M. Harsh and S. S. Stevens, "A Mechanical Correlator," *American Journal of Psychology*, 51: 727–30; October 1938.

$$P_{xy} = \tfrac{1}{2}(X + Y) - D$$

giving

$$\frac{\tfrac{1}{2}(X + Y) - D}{\sqrt{XY}} \tag{2}$$

r is thus computed from three moments of inertia of the system; σ_x and σ_y are each computed from one of these same moments (if I is known beforehand). This formula for r can be simplified for ordinary computation by reexpressing it in the rapidly convergent power series

$$= \left[1 - \frac{D}{\tfrac{1}{2}(X + Y)}\right]\left[1 + \frac{1}{2}\left(\frac{X - Y}{X + Y}\right)^2 + \frac{3}{8}\left(\frac{X - Y}{X + Y}\right)^4 + \cdots\right]. \tag{2'}$$

To find r from this formula, first find

$$d = |X - Y|,$$
$$a = \tfrac{1}{2}(X + Y),$$

and d/a.
If $d/a < .4$,

$$r = r_0 = 1.00 - \frac{D}{a} \tag{2a}$$

with an error less than .02 from the approximation.
If $d/a < .8$,

$$r = r_0 - \frac{r_0}{8}\left(\frac{d}{a}\right)^2 \tag{2b}$$

with an error less than .01 from the approximation. The condition for using equation 2b means that neither of the moments X and Y may be more than 2.2 times the other; to avoid such a case, the geometrical ranges of x and y on the web (which vary approximately with X and Y) can be adjusted beforehand to approximate equality by giving each coordinate an appropriate scale factor in plotting. Higher precision than equation 2b is probably not worth while, since the statistical probable error of r is greater than .01 unless the system has more than 500 cases and r is greater than .90.

MECHANICAL ARRANGEMENT

The physical apparatus consists of the weighted "points," the wire mesh or "graph," and a torsion pendulum for measuring the three moments of inertia needed in the above formulae. The "points" are brass escutcheon pins ¼ inch long, which can be bought at a hardware store. (Lead pins would be better, if obtainable.) The "graph" is a flat piece of window screen (iron will do, but aluminum is better) of 12 by 12 mesh, cut square with 51 meshes (or any other convenient odd number) to a side; the two center rows of meshes (x and y axes) and every tenth row in each direction are marked lightly with red paint to help in locating positions on the graph. The moments of inertia of the mesh alone about the three axes used in the formulae, X_0, Y_0, D_0, are constants and need be measured only once; they must be subtracted from the moments measured in a problem to give the moments due to the escutcheon pins alone. If the mesh is accurately square,

$$X_0 = Y_0 = D_0.$$

The torsion pendulum is a 10 inch length of .009 inch steel music wire, hung vertically from a solder joint on a stable metal support, and having soldered at its lower end a small tin clip to grasp the wire mesh. When the clip grips the edge of the mesh, the latter hangs in a vertical plane; if it is then rotated 180° about the axis of the wire and released, it will oscillate back and forth for 20 or more cycles (depending on the distribution of pins and dimensions of the mesh and music wire) until air resistance brings it to rest. The period, T, of the oscillation is related to the moment of inertia about the axis of rotation as follows:

$$T_{y_0}^2 = kY_0, \quad \text{etc.},$$

where k is a constant. Since in our formulae only ratios of moments are involved, we can measure the moments in arbitrary units; in fact we can replace each moment in our formulae by the corresponding value of t^2, where t is the time for the pendulum to make any convenient constant number of oscillations, say 25.

The times are determined by a stopwatch reading to 0.1 second; the usual precautions in counting oscillations must be observed. For ease of counting oscillations, the length of the music wire should be set so that a single vibration takes between 1 and 2 seconds; for accuracy, the number of oscillations to be counted should be chosen to give times of at least 30 seconds with the pins located in the mesh for an average problem. When graphing less than 35 sets of values, it may be convenient to double all the moments and increase the times by putting two pins side by side on the graph to represent each point.

PROCEDURE

Then to solve a problem, we first need to know the constants of the apparatus, X_0, Y_0, D_0, and I:

$$X_0 = t_{x_0}^2$$

$$Y_0 = t_{y_0}^2$$

$$D_0 = t_{d_0}^2$$

$$I = \frac{t_1^2 - t_{y_0}^2}{20,000}$$

(t_1 is the time about the y-axis of the mesh when it has 50 pins in it, 25 on the line $x=20$ and 25 on the line $x=-20$.)

Next from the given matched set of x and y-values we find by arithmetic the mean of each set. These means, rounded off to the nearest unit, are then the values of the x and y-axes on the wire mesh. (See the appendix for a discussion of the error introduced in rounding-off.) A scale is chosen so that the ranges of x and y approximately fill the mesh in each direction. The pins are pressed into the mesh as shown in the Chart; this operation goes more easily if the mesh is supported at the edges a little way above the work table. An aid to locating values on the graph quickly is a sheet of paper under the mesh marked with the numerical values assigned to the red-painted rows and columns. The time necessary for graphing is comparable to that of graphing the same points with a pencil and cross-section paper; our average is about 15 minutes to insert and check 50 points.

When the graph is completed, the mesh is clipped to the music wire at the positive x-axis, and 25 oscillations timed, giving t_x; then the same for the positive y-axis, giving t_y; then the same for the 45° angle (the clip attached to the upper right-hand corner of the mesh), giving t_d. (See Chart.)

Now we compute

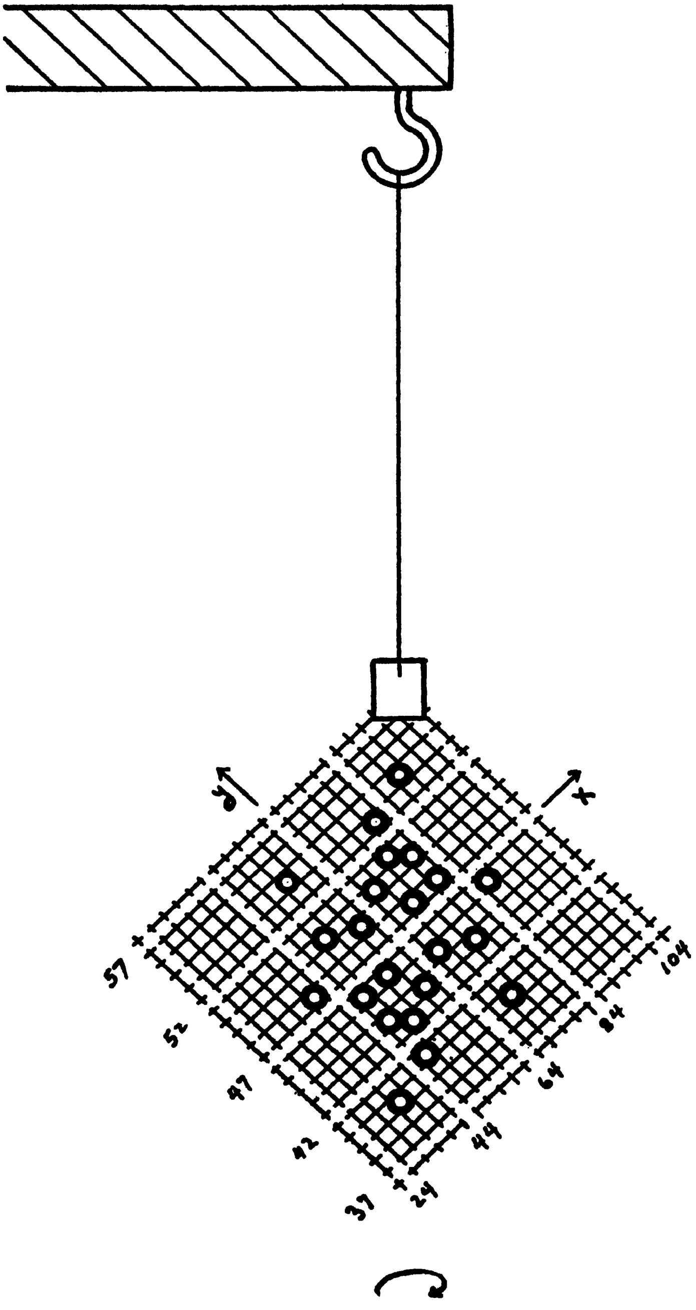
$$X = t_x^2 - X_0,$$

$$Y = t_y^2 - Y_0,$$

$$D = t_d^2 - D_0;$$

the standard deviations and correlation coefficient are obtained immediately on substituting these into equations 1 and 2. Note that σ from equation 1 is in mesh units and must be multiplied by the proper scale factor to put it in the numerical units of the problem.

If the value of D is greater than a, a negative correlation is indicated;


y
x

this can be determined more accurately by using in equation 2a the quantity D' instead of D, and reversing the sign of r, D' being found by moments about a diagonal axis through the upper *left*-hand corner of the mesh.

ERRORS AND CORRECTIONS

With t_x and t_y about 30 seconds and t_{x_0} and t_{y_0} about 15 seconds, an experimental probable error of 0.1 second leads to a probable error in each standard deviation of about 0.6 per cent. These same values lead to an experimental error of 0.8 per cent in the D/a term of equation 2a; this produces in r an experimental uncertainty changing linearly from .008 when r is zero to .000 when r is unity. If a probable error of 0.1 second is assigned to t_d and to t_{d_0} in the same problem the r acquires an additional experimental error of .011 at zero changing to .007 at unity. This suggests that t_d should be determined twice for increased accuracy, while one measurement is enough for t_x and t_y.

The escutcheon pins have so far been treated as mass points. This is a good approximation unless the highest accuracy is desired; then it is necessary to consider the effect of the finite length of the pins. This effect increases the apparent value of each of the three moments by a constant amount, NI_z, which can be found as shown in the appendix; this quantity must then be subtracted from X, Y, and D as determined above before using them in the formulae. In the problems which have been worked out so far, the effect of this correction on the computed r-value has been about .01.

Several correlation coefficients have been determined on the apparatus described here, and the agreement in each case with the results of arithmetical computation by standard methods has been satisfactory.

The device can also be used on problems other than that of bivariate correlation, in particular for measurement of the standard deviation, σ, or of the sum of the squares of the deviations, $N\sigma^2$, of a single set of values, and for measurement of the product-moment, P_{xy}, of two sets of values, a quantity needed for many statistical problems of all kinds. However, such applications will be immediately apparent from the discussion above without further elaboration.

We hope that the somewhat detailed treatment of design, operation, and sources of error given here will enable others to construct and use this apparatus. The dimensions and specific values mentioned are intended as guides to such construction and use and not as any limitation on the application of the physical principles involved.

APPENDIX

The added moment, I_s, due to the finite length of each pin, can be found by measuring the moment I_a about the y-axis with 50 pins 20 units from that axis, and then the moment I_b with 50 pins only 2 units from that axis.

$$I_a = I\ (\text{ideal})_{50,20} + 50I_s$$
$$I_b = I\ (\text{ideal})_{50,2} + 50I_s.$$

But

$$I\ (\text{ideal})_{50,2} = .01I\ (\text{ideal})_{50,20} = .01(I_a - 50I_s).$$

So

$$I_s = \frac{1}{49.5}(I_b - .01I_a) = \frac{1}{49.5}\left[t_b{}^2 - t_{y_0}{}^2 - .01(t_1{}^2 - t_{y_0}{}^2)\right].$$

An error is also introduced in rounding off the mean to the nearest unit before plotting. This error is due to the fact that the centers of gravity of the mesh and the points do not coincide, and the system therefore does not oscillate about the center of gravity of the points as we have assumed. However, if the points are, say, three times as heavy as the mesh, the center of gravity of the points must lie within 0.2 mesh units of the axis of oscillation, and the increase in moment is at most

$$N(.2)^2 + \tfrac{1}{3}N(.5)^2 = .12N$$

with the pin mass and the mesh width taken as unity. In the same units, the correction due to finite pin length is about

$$N(1.0)^2$$

since the radius of gyration of each pin ($\frac{1}{4}$ inch long) is about one mesh unit (1/12 inch). The effect of the rounding-off is therefore considerably smaller even than the quantity I_s and may be ignored.

THE EVIDENCE FOR PERIODICITY IN SHORT TIME SERIES

BY TRUMAN L. KELLEY
Harvard University

MANY TIME SERIES are of short duration because the recording of the event has, of necessity, proceeded for but a short time. Though it may be necessary that the student bide his time and wait for the years to pass in order to secure sufficient data, we should come to this conclusion only after having exhausted the possibilities of small sample theory. A method which applies to short series and which enables a judgment of the hypothesis with fiducial limits is important in that, whatever the time span, the issue can be investigated and the hypothesis substantiated, or rejected, within the fiducial limits set by the investigator.

Let the time variable be X_1 and the function in which we postulate periodicity be X_0. We shall employ M for the mean, V for the variance, $\bar{X}_0$ to indicate an X_0 as estimated from X_1 by means of a linear regression line, $\bar{\bar{X}}_0$ as estimated by means of a quadric regression line, $\bar{\bar{\bar{X}}}_0$ as estimated by means of a cubic regression line, etc. As is usual in multiple correlation, the dot in a subscript will indicate "independent of," thus $X_{0.1}$ is that part of X_0 which is independent of X_1, $X_{0.11^2}$ is that part of X_0 which is independent of X_1, X_1^2, and $X_{0.11^21^3}$ is that part of X_0 which is independent of X_1, X_1^2 and X_1^3, etc.

Before investigating periodicity we should first take out any trend that may exist. For the purposes of precise test, as hereinafter applied, it is important that the trend be represented by a line that places linear restrictions only upon X_0. Consider the following series of estimates of X_0:

$$\text{Estimate of } X_0 = M_0 \qquad X_1^0 \text{ regression} \tag{1}$$

$$\bar{X}_0 = a + bX_1 \qquad \text{linear regression} \tag{2}$$

$$\bar{\bar{X}}_0 = A + BX_1 + CX_1^2 \qquad \text{quadric regression} \tag{3}$$

$$\bar{\bar{\bar{X}}}_0 = \alpha + \beta X_1 + \gamma X_1^2 + \delta X_1^3 \qquad \text{cubic regression.} \tag{4}$$

When no trend is removed, equation (1) gives the estimate of X_0 and the quantities related to X_1 in which we seek periodicity are $(X_0 - M_0)$ of the following equation, which expresses X_0 as the sum of independent parts:

$$X_0 = M_0 + (X_0 - M_0). \tag{5}$$

If the number of observations in the series is N, the corresponding degrees of freedom equation is

$$(N-1) = 0 + (N-1) \tag{6}$$

and the corresponding equation of variances is $V_0 = 0 + V_0$. (7)

When a linear trend is removed by equation (2), we have

$$X_0 = \bar{X}_0 + X_{0.1} \tag{8}$$

$$(N-1) = 1 + (N-2) \tag{9}$$

$$V_0 = V(\bar{X}_0) + V_{0.1}. \tag{10}$$

To ascertain if the linear trend is greater than chance may be expected to yield, we compute the variance ratio $F_{1(N-2)}$ having one degree of freedom in the numerator and $(N-2)$ degrees of freedom in the denominator:

$$F_{1(N-2)} = \frac{V(\bar{X}_0)/1}{V_{0.1}/(N-2)} = \frac{V(\bar{X}_0)(N-2)}{V_{0.1}(1)}. \tag{11}$$

P from this F gives the probability that a divergence from (1), or zero regression, as great as that observed, b of equation (2), would arise as a matter of chance if the true value of b is zero. Values of P progressively less and less than .5 provide greater and greater evidence that b is not a chance deviation from zero.

Having, by this procedure, established that there is a trend, we can next investigate to see if it is of a quadric curvilinear sort. By the method of least squares we obtain the quadric regression equation (3) and related constants, and write:

$$X_0 = \bar{\bar{X}}_0 + X_{0.11^2} \tag{12}$$

$$(N-1) = 2 + (N-3) \tag{13}$$

$$V_0 = V(\bar{\bar{X}}_0) + V_{0.11^2}. \tag{14}$$

Also X_0 may be divided into independent parts thus:

$$X_0 = (\bar{\bar{X}}_0 - \bar{X}_0) + \bar{X}_0 + X_{0.11^2} \tag{15}$$

$$(N-1) = 1 + 1 + (N-3) \tag{16}$$

$$V_0 = [V(\bar{\bar{X}}_0) - V(\bar{X}_0)] + V(\bar{X}_0) + V_{0.11^2}. \tag{17}$$

The variance ratio

$$F_{1(N-3)} = \frac{[V(\bar{\bar{X}}_0) - V(\bar{X}_0)](N-3)}{V_{0.11^2}} \tag{18}$$

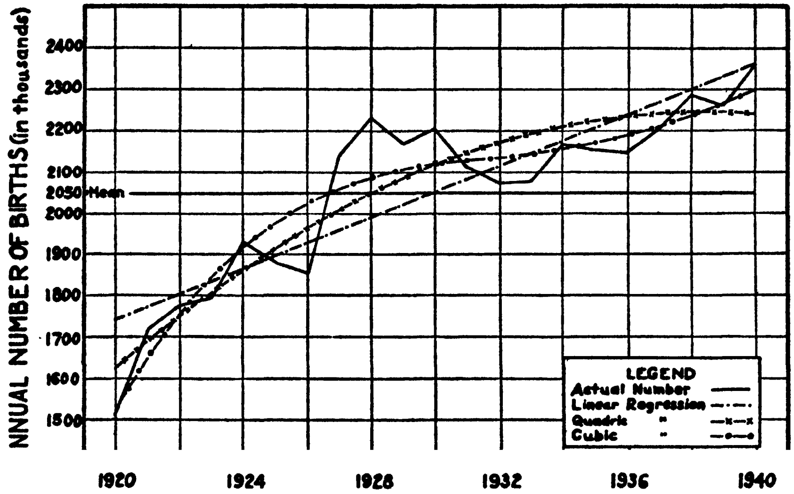
having one degree of freedom in the numerator and $(N-3)$ in the denominator, will yield a P that tests the significance of C of equation (3).

If this is significant then the trend to be removed is that given by the quadric regression line.

A similar procedure can test for cubic regression, δ of equation (4), but should hardly be resorted to if the interest is in the existence of periodicity, for a cubic will itself be a close approximation to a complete cycle of a certain sine curve. The cubic curve of Chart I, which cuts the quadric much as would a sine curve with a period of 16 years, is an illustration of this.

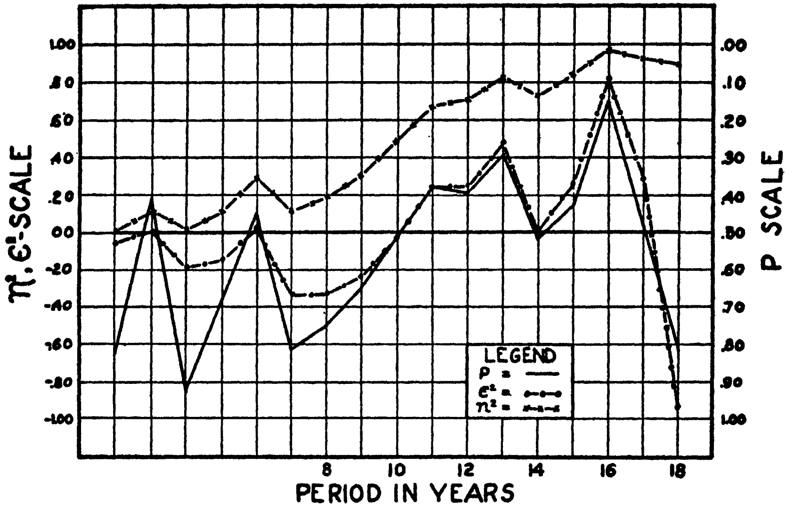
CHART I

BIRTHS IN U. S. 1920–1940



Let us say that we have removed a quadric trend, obtaining residuals $X_{0.11}'$, which are to be investigated for periodicity. We shall use a periodogram tehnique with two modifications. The usual periodogram is as shown in the dot line of Chart II. The correlation ratio squared, η^2, is plotted for successively greater periods, T. The argument is that in the neighborhood of that period yielding the largest η^2 will be found the actual period of the data. For the argument to be precise it is necessary that a probability statement be attached to the outcome. Thus the essential modification recommended is that instead of an ordinate η^2, we employ the ordinate $(1-P)$, the P being that from a variance ratio, the numerator variance being that of the means of the classes into which the data have been grouped when items of the same phase for the period T constitute a group, and the denominator variance is the customary residual, or error, variance. The second modification gives an approximation to this first while at the same time it gives a quantitative statement of the square of the correlation between time treated as a periodical variable of period T and the function, $X_{0.11^2}$, being investigated. This second modification is to plot ϵ^2, the correlation ratio squared corrected for fineness of grouping,[1] instead of $(1-P)$, with which it is highly correlated, and instead of the usual η^2, from which it diverges considerably because of the systematic error inherent in η^2. The illustration which follows will elucidate the relationships between $(1-P)$, ϵ^2 and η^2.

CHART II
PERIODOGRAM



Let there be N equally time-spaced observations X_0. Let a quadric trend be removed, giving N residuals $X_{0.11^2}$ having $(N-3)$ degrees of freedom. Let the period under investigation be T of the equally spaced time intervals. Of necessity $1<T<(N-2)$ and the practical limits of utility of T, fractional or integral, may be taken such that $2 \leqq T \leqq (N-3)$. If a linear, instead of quadric, trend is removed $2 \leqq T \leqq (N-2)$, and if a zero trend, equation (1), is involved then $2 \leqq T \leqq (N-1)$. If N is exactly divisible by T, there are T classes each having N/T measures, $X_{0.11^2}$, of the same phase in it. If N is not exactly divisible by T, we have a variable number of measures from phase to phase and the number of classes will be some number, which we will call k, greater than N/T. Let the means of the $X_{0.11^2}$ values in these classes be M_{0a}, M_{0b}, $\cdots$, M_{0k}. The variance of these means, appropriately weighted with the numbers of cases in each class, is designated $V(M_{0K})$. The mean for each class is independent of that for every other class except that the mean of the means is equal to zero, so that there are just $(k-1)$ degrees of freedom. Letting i take all values from a to k, we can express the variable $X_{0.11^2}$ as equal to the sum of two independent parts,

---

[1] Truman L. Kelley, "An Unbiased Correlation Ratio Measure," *Proceedings of the National Academy of Sciences*, Vol. 21, No. 9, September 1935, pp. 554–559. Also discussed by C. C. Peters and W. VanVoorhis, *Statistical Procedures and their Mathematical Bases*, 1940, Ch. XI and XII.

$$X_{0.11^2} = M_{0i} + X_{0.11^2 i} \tag{19}$$

$$(N-3) = (k-1) + (N-2-k) \tag{20}$$

$$V_{0.11^2} = V(M_{0K}) + V_{0.11^2 i} \tag{21}$$

$$F_{(k-1)(N-2-k)} = \frac{V(M_{0K})/(k-1)}{V_{0.11^1 i}/(N-2-k)} = \frac{[V(M_{0K})](N-2-k)}{[V_{0.11^2} - V(M_{0K})](k-1)}. \tag{22}$$

As P from this variance ratio is less and less than .5, there is greater and greater evidence that a period in the neighborhood of period T exists. The usual procedures for testing periods which are fractional introduces no complication if proper allowance is made for the decrease in N because certain X_0 values are not used.

The computations are quite simple, even that of P from F, if Pearson's *Tables of the Incomplete Beta Function* (1934) are used. The illustration following suggests that ϵ^2 may be used in lieu of $(1-P)$. It also has the advantage of being an unbiased measure of the strength of the correlation maintaining. Chart II shows the relationship of $(1-P)$, ϵ^2, and η^2, for a certain problem.

The data chosen to illustrate the procedure described are the number of births in the United States for the years 1920 to 1940. The series is short, $N=21$. A trend of uncertain nature is present. The likelihood of periodicity is sufficiently small to lead to borderline issues and critical tests. The Table herewith gives the basic data and certain derived data used for testing for different periods.

The columns in order provide:

Column 1: Year
Column 2: Number of births
Column 3: X_0, number estimated by (23)
Column 4: $X_{0.11^2}$, residual deviations to be tested for periodicity.
Column 5: Column 4 values opposite a are in the same phase and thus constitute one class. Values opposite b are in the same, and a different, phase and constitute the second class, when the period being investigated is 2 years.

Column 6, 7, . . . , 21: are similar for periods of 3, 4, . . . , 18 years.

At the feet of columns 5, 6, . . . , 21 are recorded, in order, BIRTHS IN THE UNITED STATES 1920-1940

X_1 Year	X_0 Births	$\frac{\overline{\overline{X}}_0}{100}$	$\frac{X_{0,u^2}}{100}$	Phases for periods T																
				2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1920	1508874	16219	−1130	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*
1	1714261	16885	258	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*
2	1774911	17514	235	*a*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*
3	1792646	18105	− 179	*b*	*a*	*d*	*d*	*d*	*d*	*d*	*d*	*d*	*d*	*d*	*d*	*d*	*d*	*d*	*d*	*d*
4	1930614	18658	648	*a*	*b*	*a*	*e*	*e*	*e*	*e*	*e*	*e*	*e*	*e*	*e*	*e*	*e*	*e*	*e*	*e*
5	1878880	19174	− 385	*b*	*c*	*b*	*a*	*f*	*f*	*f*	*f*	*f*	*f*	*f*	*f*	*f*	*f*	*f*	*f*	*f*
6	1856068	19652	−1091	*a*	*a*	*c*	*b*	*a*	*g*	*g*	*g*	*g*	*g*	*g*	*g*	*g*	*g*	*g*	*g*	*g*
7	2137836	20093	1285	*b*	*b*	*d*	*c*	*b*	*a*	*h*	*h*	*h*	*h*	*h*	*h*	*h*	*h*	*h*	*h*	*h*
8	2233149	20496	1835	*a*	*c*	*a*	*d*	*c*	*b*	*a*	*i*	*i*	*i*	*i*	*i*	*i*	*i*	*i*	*i*	*i*
9	2169920	20862	837	*b*	*a*	*b*	*e*	*d*	*c*	*b*	*a*	*j*	*j*	*j*	*j*	*j*	*j*	*j*	*j*	*j*
1930	2203958	21190	850	*a*	*b*	*c*	*a*	*e*	*d*	*c*	*b*	*a*	*k*	*k*	*k*	*k*	*k*	*k*	*k*	*k*
1	2112760	21480	− 352	*b*	*c*	*d*	*b*	*f*	*e*	*d*	*c*	*b*	*a*	*l*	*l*	*l*	*l*	*l*	*l*	*l*
2	2074042	21733	− 993	*a*	*a*	*a*	*c*	*a*	*f*	*e*	*d*	*c*	*b*	*a*	*m*	*m*	*m*	*m*	*m*	*m*
3	2081232	21949	−1137	*b*	*b*	*b*	*d*	*b*	*g*	*f*	*e*	*d*	*c*	*b*	*a*	*n*	*n*	*n*	*n*	*n*
4	2167636	22127	− 451	*a*	*c*	*c*	*e*	*c*	*a*	*g*	*f*	*e*	*d*	*c*	*b*	*a*	*o*	*o*	*o*	*o*
5	2155105	22267	− 716	*b*	*a*	*d*	*a*	*d*	*b*	*h*	*g*	*f*	*e*	*d*	*c*	*b*	*a*	*p*	*p*	*p*
6	2144790	22370	− 922	*a*	*b*	*a*	*b*	*e*	*c*	*a*	*h*	*g*	*f*	*e*	*d*	*c*	*b*	*a*	*q*	*q*
7	2203337	22435	− 402	*b*	*c*	*b*	*c*	*f*	*d*	*b*	*i*	*h*	*g*	*f*	*e*	*d*	*c*	*b*	*a*	*r*
8	2286962	22462	408	*a*	*a*	*c*	*d*	*a*	*e*	*c*	*a*	*i*	*h*	*g*	*f*	*e*	*d*	*c*	*b*	*a*
9	2265558	22453	203	*b*	*b*	*d*	*e*	*b*	*f*	*d*	*b*	*j*	*i*	*h*	*g*	*f*	*e*	*d*	*c*	*b*
1940	2360399	22405	1199	*a*	*c*	*a*	*a*	*c*	*g*	*e*	*c*	*a*	*j*	*i*	*h*	*g*	*f*	*e*	*d*	*c*
			$\eta^2=$	.004	.115	.014	.112	.298	.109	.186	.312	.489	.665	.708	.826	.725	.837	.969	.920	.898
			$F_{(k-1)(N-2-k)}=$	.073	1.042	.414	.439	1.104	.244	.360	.568	.958	1.590	1.541	2.368	1.016	1.468	6.240	1.130	.490
			$P=$	.825	.408	.923	.682	.449	.817	.750	.649	.512	.378	.391	.291	.519	.429	.152	.485	.814
			$\epsilon^2=$	−.054	.005	−.183	−.142	.028	−.337	−.331	−.238	−.022	.248	.248	.477	.012	.267	.814	.283	−.927

Source: *World Almanac, 1943*, p. 472.

η^2, $F_{(k-1)(N-2-k)}$, P corresponding, and ϵ^2.

We first investigate trend. X_0 = number of births per year. X_1 = date and for convenience we let $x_1 = X_1 - 1930$. We find:

$$M_0 = 2050140, \quad \text{and} \quad V_0 = 4617948 \times 10^4$$
$$\overline{X}_0 = 2050140 + 30929x_1$$
$$r_{01}^2 = .75956.$$

To test the significance of the deviation of 30929 from zero we compute

$$F_{1,19} = \frac{V_0 r_{01}^2 (N-2)}{V_0(1 - r_{01}^2)} = 60.022$$

which yields a very small P. We next fit a quadric regression line, obtaining

$$\overline{\overline{X}}_0 = 2118995 + 30929x_1 - 1877.9x_1^2 \qquad (23)$$
$$r^2_{0.11^2} = .84113$$
$$F_{1,18} = \frac{(V_0 r^2_{0.11^2} - V_0 r_{01}^2)(N-3)}{V_0(1 - r^2_{0.11^2})} = 9.2419$$

yielding $P = .048$. Having odds of 952 to 48 that -1877.9 is not a chance deviation from zero we take (23) as the trend line and compute residuals

$$X_{0.11^2} = X_0 - \overline{\overline{X}}_0$$

as recorded in column 4 of the Table herewith.

We have computed a cubic regression line, not as a part of the procedure of determining periodicity, but because it is interesting to compare such a line with the evidence available as to periodicity.

$$\overline{\overline{\overline{X}}}_0 = 2118994 + 11149x_1 - 1877.8x_1^2 + 300.61x_1^3$$
$$r^2_{0.11^21^3} = .89917$$
$$F_{1,17} = \frac{(V_0 r^2_{0.11^21^3} - V_0 r^2_{0.11^2})(N-4)}{V_0(1 - r^2_{0.11^21^3})} = 9.7856$$

yielding $P = .045$. This establishes the non-chance nature of the coefficient 300.61. We will later note that we do not establish periodicity with the certainty that we here establish the fitness of a cubic regression line.

To test for the existence of a two-year period the residuals of column four are to be grouped into two classes, *a* and *b* as indicated in column five. The *a* class has 11 measures and yields a mean, M_{0a}, and the 10 measures of class b yield a mean, M_{0b}. Weighting these means 11 and 10 respectively and computing their variance, we obtain $V(M_{02})$, the subscript 2 being k, the number of classes involved. Analyzing variance as indicated in (21) and employing (22) we have: $F_{1,17}=.073$, yielding $P=.825$, which, being $>.5$, yields no evidence whatever that a period of two years exists. Similar computations for $T=3$, $T=4$, $\cdots$, $T=18$ have been made and are as recorded at the feet of the appropriate columns of the table. The smallest P is .152, found when $T=16$. Thus the odds are about five to one that a period in the neighborhood of 16 years exists. The F for this case is an $F_{15,3}$. We have but three degrees of freedom available in the error variance. Should a real period of 16 years exist it is not surprising that data covering but 21 years is unable to establish it with satisfactory certainty.

The advantages of employing $(1-P)$ in lieu of η^2, in plotting and interpreting a periodogram wherein the number of degrees of freesdom is small, follow from the argument to this point and are exemplified in Chart II.

To a degree the advantages of $(1-P)$ are also inherent in ϵ^2, as is indicated by the periodogram curve based upon ϵ^2. The earlier derivation by the writer of ϵ^2 gave the relationship

$$\epsilon^2 = 1 - \frac{N-1}{N-k}(1-\eta^2) \tag{24}$$

but this was postulated upon using a zero order regression equation, (1). When a linear trend is taken out the relationship is

$$\epsilon^2 = 1 - \frac{N-2}{N-1-k}(1-\eta^2). \tag{25}$$

If a quadric trend is removed we have

$$\epsilon^2 = 1 - \frac{N-3}{N-2-k}(1-\eta^2) \tag{26}$$

and so forth for the removal of trends of higher parabolic order. The values of ϵ^2 given in the table were computed by (26). The median ϵ^2 value of the table is .012 which differs but slightly from .000, the value to be expected if there is no period in the data.

The periodogram analysis here discussed may be thought of as terminating the problem only in case no period is discovered. If the analysis does establish one or more periods, their further specification could be accomplished by fitting, say, a Fourier series, by a method which is appropriate when the time span of the data is not an exact multiple of the period, or periods.

CORRELATION CONCEPTS AND THE DOOLITTLE METHOD

By Dudley J. Cowden
University of North Carolina

Among the many concepts of correlation, one which is especially well adapted to the Doolittle method, outlined by Paul S. Dwyer in the December, 1942, issue of this Journal, is that the coefficient of determination r^2 is the proportion of variation in the dependent variable that has been explained by variations in the independent variable or variables; or it is the ratio of the explained variation to the total variation. The chief purpose of this article is to indicate how the Doolittle method fits in with this concept.

In this paper we shall use $x=X-\overline{X}$ to indicate a deviation from the mean, and following Dwyer, the highest subscript will be assigned to the dependent variable. Where two subscripts follow the letter x, the first one will refer to the dependent variable, and the meaning of the symbol will be a deviation from the mean of the explained portion of the dependent variable, i.e., the value of a deviation from the mean as computed from the estimating equation involving the independent variable indicated. If there are two or more independent variables in the equation, the subscripts referring to those variables will be enclosed in parentheses. A subscript or subscripts following a decimal point will always indicate that the variable, or those variables, have been "held constant." For a three variable problem, the various coefficients of determination involving the dependent variable, and their conceptual formulas are given below.

Simple correlation

$$r^2_{31} = \frac{\Sigma x^2_{31}}{\Sigma x^2_3} \qquad\qquad r^2_{32} \quad \begin{matrix}\Sigma x^2_{32}\\ \Sigma x^2_3\end{matrix}$$

Multiple correlation

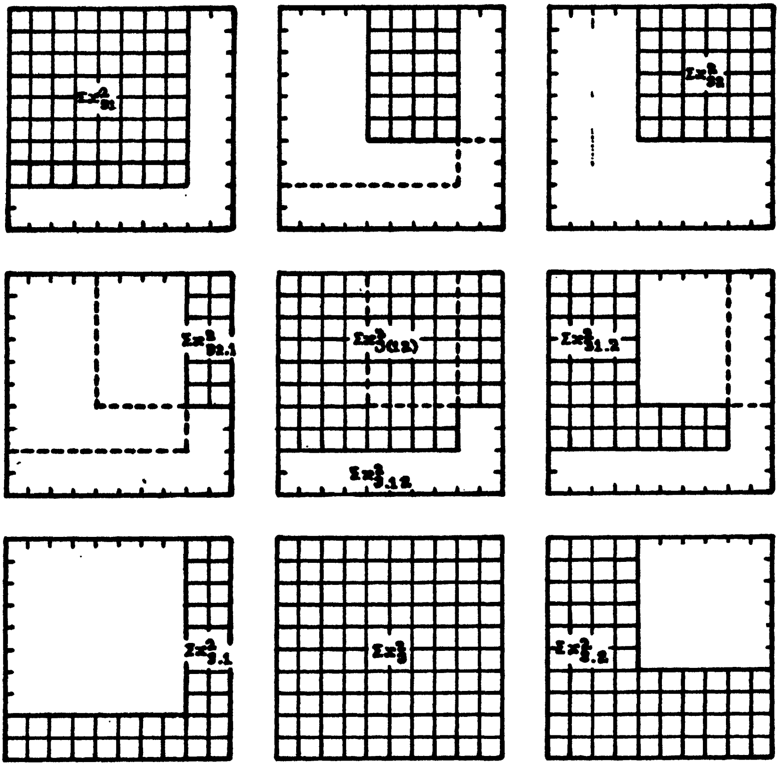
$$r^2_{3(12)} = \frac{\Sigma x^2_{3(12)}}{\Sigma x^2_3}$$

Partial correlation

$$r^2_{32.1} = \frac{\Sigma x^2_{32.1}}{\Sigma x^2_{3.1}} \qquad\qquad r^2_{31.2} = \begin{matrix}\Sigma x^2_{31.2}\\ \Sigma x^2_{3.2}\end{matrix}$$

Reference to the accompanying diagram serves to visualize the concept. The total variation Σx^2_3 is represented by the middle square in the bottom row. In the two upper corners are shown the gross amounts of variation Σx^2_{31} and Σx^2_{32} explained by the two independent variables. Because of factors common to these two variables there is an overlapping in the explained areas, which is indicated by the middle rectangle at the top of the diagram. In the central square of the diagram $\Sigma x^2_{3(12)}$ represents the amount of variation explained by the two independent variables taken together. To the left of that polygon is shown the additional amount of variation, $\Sigma x^2_{32.1}$ accounted for by variable 2 over and above the gross amount explained by variable 1; while to the right is shown the net variation $\Sigma x^2_{31.2}$ explained by variable 1. The diagram indicates that the amount of variation explained by variables 1 and 2 taken together is made up of:

DIAGRAMMATIC REPRESENTATION OF TYPES OF VARIATION INVOLVED IN SIMPLE, MULTIPLE, AND PARTIAL CORRELATION



(a) gross explained by 1 plus net explained by 2; or
(b) gross explained by 2 plus net explained by 1; or
(c) net explained by 1 plus net explained by 2 plus amount explained by factors common to 1 and 2.

The amount of unexplained variation remaining after using variable 1 alone $\Sigma x^2_{3.1}$, and therefore the amount which it is sought to explain by the net contribution of variable 2, is shown in the lower left hand corner of the diagram. Finally, the amount still to be explained after variable 2 alone has been used $\Sigma x^2_{3.2}$ is shown in the lower right hand corner. These two amounts are the denominators for the partial correlation coefficients $r^2_{32.1}$ and $r^2_{31.2}$ respectively.

Since each of the nine large squares is composed of 100 small squares it is easy to compute the area of the different polygons. If this is done it will be seen that the values of the various coefficients of determination are as follows:

$$r^2_{31} = \frac{64}{100} = .64 \qquad r^2_{32} = \frac{36}{100} = .36$$

$$r^2_{3(12)} = \frac{76}{100} = .76$$

$$r^2_{32.1} = \frac{12}{36} = .333 \qquad r^2_{31.2} = \frac{40}{64} = .625.$$

We shall now turn our attention to computational methods, using four variables as an illustration. First, to make estimates of variable 4 from variable 1, we may use the estimating equation $x_{41} = b_{41}x_1$, in which $b_{41} = \dfrac{\Sigma x_4 x_1}{\Sigma x^2_1}\cdot$ We could then obtain Σx^2_{41} by squaring the N values of x_{41} and summing them. The Doolittle solution accomplishes the same result by the expression $\Sigma x^2_{41} = b_{41}\Sigma x_4 x_1$.

The amount of variation in variable 4 explained by variable 2, with variable 1 held constant, could literally be obtained by first adjusting variables 4 and 2 for variable 1 and then proceeding in a manner similar to that in the preceding paragraph. Thus, we could obtain $\Sigma x^2_{42.1}$ as follows:

(1) Compute the N values of $x_{4.1} = x_4 - x_{41}$, and the N values of $x_{2.1} = x_2 - x_{21}$;

(2) Compute the N values of $x_{4.1}x_{2.1}$ and sum them;

(3) Square the N values of $x_{2.1}$ and sum them;

(4) Compute $b_{42.1}=\dfrac{\Sigma x_{4.1}x_{2.1}}{\Sigma x^2_{2.1}}$;

(5) Compute $\Sigma x^2_{42.1}=b_{42.1}\Sigma x_{4.1}x_{2.1}$.

Another method, which serves further to clarify our concept of correlation, is based on the principles:

> Explained variation plus unexplained variation equals total variation;
> Explained covariation plus unexplained covariation equals total covariation.

By the first principle, as is well known,

$$\Sigma x^2_{2.1} = \Sigma x^2_2 - \Sigma x^2_{21};$$

and by the second,

$$\Sigma x_{4.1}x_{2.1} = \Sigma x_4x_2 - \Sigma x_{41}x_{21}.$$

Thus we could compute the N values of x_{41} and the N values of x_{21}, take the product of each of the N pairs, sum them, and subtract the sum from Σx_4x_2. However, just as

$$\begin{aligned} \Sigma x^2_{21} &= b_{21}\Sigma x_2x_1, \quad \text{and} \\ \Sigma x^2_{41} &= b_{41}\Sigma x_4x_1; \quad \text{so} \\ \Sigma x_{41}x_{21} &= b_{21}\Sigma x_4x_1, \quad \text{or} \\ & b_{41}\Sigma x_2x_1. \end{aligned}$$

The Doolittle solution employs whichever of these two methods for $\Sigma x_{41}x_{21}$ that the computer prefers.

By extension of the method indicated above, we obtain, by the Doolittle solution:

$$\begin{aligned} \Sigma x_{4.12}x_{3.12} &= \Sigma x_4x_3 - \Sigma x_{4(12)}x_{3(12)} \\ &= \Sigma x_4x_3 - \Sigma x_{41}x_{31} - \Sigma x_{42.1}x_{32.1} \\ &= \Sigma x_4x_3 - b_{31}\Sigma x_4x_1 - b_{32.1}\Sigma x_{4.1}x_{2.1}, \quad \text{or} \\ & \Sigma x_4x_3 - b_{41}\Sigma x_3x_1 - b_{42.1}\Sigma x_{3.1}x_{2.1}; \\ b_{43.12} &= \frac{\Sigma x_{4.12}x_{3.12}}{\Sigma x^2_{3.12}}; \\ \Sigma x^2_{43.12} &= b_{43.12}\Sigma x_{4.12}x_{3.12}. \end{aligned}$$

The total amount of variation explained by several variables is obtained by adding to the gross amount explained by variable 1 the successive additional amounts explained by other variables as they are included successively in the estimating equation. Thus:

$$\Sigma x^2_{4(123)} = \Sigma x^2_{41} + \Sigma x^2_{42.1} + \Sigma x^2_{43.12}.$$

The multiple coefficient of determination

$$r^2_{4(123)} = \frac{\Sigma x^2_{4(123)}}{\Sigma x^2_4}$$

and the partial coefficient of determination

$$r^2_{43.12} \quad \frac{\Sigma x^2_{43.12}}{\Sigma x^2_{4.12}} \quad \frac{\Sigma x^2_{43.12}}{\Sigma x^2_4 - \Sigma x^2_{4(12)}}$$

The Doolittle solution may also be used, though somewhat less conveniently, to obtain the above numerator by the expression

$$\Sigma x^2_{43.12} = \Sigma x^2_{4(123)} - \Sigma x^2_{4(12)}, \quad \text{where}$$
$$\Sigma x^2_{4(12)} = b_{42.1}\Sigma x_4x_1 + b_{42.1}\Sigma x_4x_2, \quad \text{and}$$
$$\Sigma x^2_{4(123)} = b_{41.23}\Sigma x_4x_1 + b_{42.13}\Sigma x_4x_2 + b_{43.12}\Sigma x_4x_3.$$

TABLE I

VALUES NEEDED FOR SIMPLE, MULTIPLE, AND PARTIAL CORRELATION AS OBTAINED BY DOOLITTLE SOLUTION

	(1)	(2)	(3)	(4)	[1]	[2]	[3]	[4]
(1)	Σx^2_1	Σx_2x_1	Σx_3x_1	Σx_4x_1	-1			
(2)		Σx^2_2	Σx_3x_2	Σx_4x_2	0	-1		
(3)			Σx^2_3	Σx_4x_3	0	0	-1	
(4)				Σx^2_4	0	0	0	-1
I	Σx^2_1	Σx_2x_1	Σx_3x_1	Σx_4x_1	-1			
I′	1	b_{21}	b_{31}	b_{41}	$-1/\Sigma x^2_1$			
II		$\Sigma x^2_{2.1}$	$\Sigma x_{3.1}x_{2.1}$	$\Sigma x_{4.1}x_{2.1}$	b_{21}	-1		
II′		1	$b_{32.1}$	$b_{42.1}$	$b_{21}/\Sigma x^2_{2.1}$	$-1/\Sigma x^2_{2.1}$		
III			$\Sigma x^2_{3.12}$	$\Sigma x_{4.12}x_{3.12}$	$b_{31.2}$	$b_{32.1}$	-1	
III′			1	$b_{43.12}$	$b_{31.2}/\Sigma x^2_{3.12}$	$b_{32.1}/\Sigma x^2_{3.12}$	$-1/\Sigma x^2_{3.12}$	
IV				$\Sigma x^2_{4.123}$	$b_{41.23}$	$b_{42.13}$	$b_{43.12}$	-1
IV′					$b_{41.23}/\Sigma x^2_{4.123}$	$b_{42.13}/\Sigma x^2_{4.123}$	$b_{43.12}/\Sigma x^2_{4.123}$	$-1/\Sigma^2_{4.123}$
i	Σx^2_1	Σx^2_{21}	Σx^2_{31}	Σx^2_{41}	$1/\Sigma x^2_1$			
ii		$\Sigma x^2_{2.1}$	$\Sigma x^2_{32.1}$	$\Sigma x^2_{42.1}$	$b^2_{21}/\Sigma x^2_{2.1}$	$1/\Sigma x^2_{2.1}$		
iii			$\Sigma x^2_{3.12}$	$\Sigma x^2_{43.12}$	$b^2_{31.2}/\Sigma x^2_{3.12}$	$b^2_{32.1}/\Sigma x^2_{3.12}$	$1/\Sigma^2_{3.12}$	
iv				$\Sigma x^2_{4.123}$	$b^2_{41.23}/\Sigma x^2_{4.123}$	$b^2_{42.13}/\Sigma x^2_{4.123}$	$b^2_{43.12}/\Sigma x^2_{4.123}$	$1/\Sigma x^2_{4.123}$
Σi	Σx^2_1	Σx^2_{21}	Σx^2_{31}	Σx^2_{41}	$1/\Sigma x^2_1$			
Σii		Σx^2_2	$\Sigma x^2_{3(12)}$	$\Sigma x^2_{4(12)}$	$1/\Sigma x^2_{1.2}$	$1/\Sigma x^2_{2.1}$		
Σii			Σx^2_3	$\Sigma x^2_{4(123)}$	$1/\Sigma x^2_{1.23}$	$1/\Sigma x^2_{2.13}$	$1/\Sigma x^2_{3.12}$	
Σiv				Σx^2_4	$1/\Sigma x^2_{1.234}$	$1/\Sigma x^2_{2.134}$	$1/\Sigma x^2_{3.124}$	$1/\Sigma x^2_{4.123}$

If the Doolittle solution is carried out as suggested here, in deviation form rather than in standard form, one is spared the labor of computing the different zero-order coefficients, as well as of converting the β coefficients into b coefficients. On the other hand, one is deprived of the preliminary information which the different zero-order coefficients provide. Also, it will often be found desirable to code the data so that the various sums of squares and sums of products do not vary too greatly in magnitude. If this is not done one may run into numbers that exceed the capacity of the calculating machine. Such coding can be done by multiplying variables, where necessary, by appropriate powers of 10; and such coding may be done after the sums of squares and sums of products have been obtained.

TABLE II

ALTERNATE NOTATIONS FOR RIGHT SIDE OF DOOLITTLE SOLUTION

	[1]	[2]	[3]	[4]
I	-1			
I′	$-1/\Sigma x^2_{1}$			
II	b_{21}	-1		
II′	$\Sigma x_2x_1/\Sigma x^2_1\Sigma x^2_{2.1}$	$-1/\Sigma x^2_{2.1}$		
III	$b_{31.2}$	$b_{32.1}$	-1	
III′	$\Sigma x_{3.2}x_{1.2}/\Sigma x^2_{1.2}\Sigma x^2_{3.12}$	$\Sigma x_{3.1}x_{2.1}/\Sigma x^2_{2.1}\Sigma x^2_{3.12}$	$-1/\Sigma x^2_{3.12}$	
IV	$b_{41.23}$	$b_{42.13}$	$b_{43.12}$	-1
IV′	$\Sigma x_{4.23}x_{1.23}/\Sigma x^2_{1.23}\Sigma x^2_{4.123}$	$\Sigma x_{4.13}x_{2.13}/\Sigma x^2_{2.13}\Sigma x^2_{4.123}$	$\Sigma x_{4.12}x_{3.12}/\Sigma x^2_{3.12}\Sigma x^2_{4.123}$	$-1/\Sigma x^2_{4.123}$
i	$1/\Sigma x^2_1$			
ii	$\Sigma x^2_{21}/\Sigma x^2_1\Sigma x^2_{2.1}$	$1/\Sigma x^2_{2.1}$		
iii	$\Sigma x^2_{31.2}/\Sigma x^2_{1.2}\Sigma x^2_{3.12}$	$\Sigma x^2_{32.1}/\Sigma x^2_{2.1}\Sigma x^2_{3.12}$	$1/\Sigma x^2_{3.12}$	
iv	$\Sigma x^2_{41.23}/\Sigma x^2_{1.23}\Sigma x^2_{4.123}$	$\Sigma x^2_{42.13}/\Sigma x^2_{2.13}\Sigma x^2_{4.123}$	$\Sigma x^2_{43.12}/\Sigma x^2_{3.12}\Sigma x^2_{4.123}$	$1/\Sigma x^2_{4.123}$
Σi	$1/\Sigma x^2_1$			
Σii	$\Sigma x^2_2/\Sigma x^2_1\Sigma x^2_{2.1}$	$1/\Sigma x^2_{2.1}$		
Σiii	$\Sigma x^2_{3.2}/\Sigma x^2_{1.2}\Sigma x^2_{3.12}$	$\Sigma x^2_{3.1}/\Sigma x^2_{2.1}\Sigma x^2_{3.12}$	$1/\Sigma x^2_{3.12}$	
Σiv	$\Sigma x^2_{4.23}/\Sigma x^2_{1.23}\Sigma x^2_{4.123}$	$\Sigma x^2_{4.13}/\Sigma x^2_{2.13}\Sigma x^2_{4.123}$	$\Sigma x^2_{4.12}/\Sigma x^2_{3.12}\Sigma x^2_{4.123}$	$1/\Sigma x^2_{4.123}$
	1			
	r^2_{21}	1		
	$r^2_{31.2}$	$r^2_{32.1}$	1	
	$r^2_{41.23}$	$r^2_{42.13}$	$r^2_{43.12}$	1

Table I gives the Doolittle solution in symbolic form. The eight sections to the table are set off by double lines. The upper left section shows the sums of squares and sums of products of deviations from the mean. In the upper right section is the identity matrix, with reversed signs. (The signs are reversed in order to obtain the correct sign for the different coefficients of estimation.) The next to the top section contains the forward solution. The next to the bottom section contains products of pairs of items in corresponding cells of the forward solution. The bottom section contains cumulative sums of the items in the same column in the section immediately above.

The left hand side of the table gives us entries by which we may easily compute, among other coefficients, the following: r^2_{41}, $r^2_{42.1}$, $r^2_{43.12}$, $r^2_{4(12)}$, $r^2_{4(123)}$. Corresponding standard errors of estimate, and variance ratios for testing significance, may likewise be computed.

The symbols recorded in the right-hand side of Table I illustrate a concept of correlation somewhat different from the concept in which we are mainly interested. A coefficient of simple or partial correlation may be thought of as the slope of the estimating coefficient adjusted for differences in variability of two variables. Thus:

$$r_{21} = b_{21}\frac{\sigma_1}{\sigma_2} = b_{21}\frac{\sigma_{1.2}}{\sigma_{2.1}}, \quad \text{and} \quad r_{41.23} = b_{41.23}\frac{\sigma_{1.234}}{\sigma_{4.123}}$$

Therefore,

$$r^2_{21} = b^2_{21}\frac{\Sigma x^2_{1.2}}{\Sigma x^2_{2.1}}, \quad \text{and} \quad r^2_{41.23} = b^2_{41.23}\frac{\Sigma x^2_{1.234}}{\Sigma x^2_{4.123}}$$

It is this method by which a new set of coefficients of partial determination are obtained from this side of the table. All that needs be done is to divide an entry in a cell in the next to the bottom section by the entry in the cell in corresponding position in the bottom section.

The derivation of the sums of products in the lower right-hand section of Table I may not be obvious; so that for $1/\Sigma x^2_{1.23}$ will be given as an illustrating procedure, using concepts developed earlier in this paper.

$$\frac{1}{\Sigma x^2_1} + \frac{b^2_{21}}{\Sigma x^2_{2.1}} = \frac{\Sigma x^2_{2.1} + b^2_{21}\Sigma x^2_1}{\Sigma x^2_1\Sigma x^2_{1.2}} = \frac{\Sigma x^2_{2.1} + b_{21}\Sigma x_2x_1}{\Sigma x^2_1\Sigma x^2_{1.2}}$$

$$= \frac{\Sigma x^2_{2.1} + \Sigma x^2_{21}}{\Sigma x^2_1\Sigma x^2_{1.2}} = \frac{\Sigma x^2_2}{\Sigma x^2_2\Sigma x^2_{1.2}} = \frac{1}{\Sigma x^2_{1.2}}$$

$$\frac{1}{\Sigma x^2_{1.2}} + \frac{b^2_{31.2}}{\Sigma x^2_{3.12}} = \frac{\Sigma x^2_{3.12} + b^2_{31.2}\Sigma x^2_{1.2}}{\Sigma x^2_{1.2}\Sigma x^2_{3.12}} = \frac{\Sigma x^2_{3.12} + \Sigma x^2_{31.2}}{\Sigma x^2_{1.2}\Sigma x^2_{3.12}}$$

$$= \frac{\Sigma x^2_{3.2}{}^*}{\Sigma x^2_{1.2}\Sigma x^2_{3.12}} = \frac{\Sigma x^2_2\Sigma x^2_{3.2}}{\Sigma x^2_2\Sigma x^2_{1.2}\Sigma x^2_{3.12}}$$

---

* See Chart for visual evidence that $\Sigma x^2_{3.12}+\Sigma x^2_{31.2}=\Sigma x^2_{3.2}$.
Algebraically we have

$$\Sigma x^2_{3.12}+\Sigma x^2_{31.2}=\Sigma x^2_3-\Sigma x^2_{3(12)}+\Sigma x^2_{3(12)}-\Sigma x^2_{32}=\Sigma x^2_3-\Sigma x^2_{32}=\Sigma x^2_{3.2}.$$

$$= \frac{\Sigma x^2_3 \Sigma x^2_{2.3}}{\Sigma x^2_3 \Sigma x^2_{2.3} \Sigma x^2_{1.23}} = \frac{1}{\Sigma x^2_{1.23}}$$

Table II shows the right-hand section of Table I (except that the identity matrix is omitted to avoid repetition and a new section is appended at the bottom), but with symbols illustrating the concept of correlation based on the ratio of explained to total variation. The derivation of these symbols and their identity with those of Table I is rather obvious. Although the entries appear somewhat more complicated, it should be noticed that the denominators of the fractions in cells of corresponding positions in the two middle sections of the table are identical. The numerators of these fractions are respectively the net explained variation and the variation remaining to be explained, which are the values needed for the various coefficients of partial determination shown in the bottom section.

OVERESTIMATION OF MEAN SQUARES BY THE METHOD OF EXPECTED NUMBERS*

BY R. E. COMSTOCK†
University of Minnesota

THE METHOD of expected numbers for analysis of the variance of data with two or more criteria of classification and disproportionate numbers in the sub-classes was introduced by Snedecor.[1] He frankly stated that the methods[2] of fitting constants and of weighted squares of means were more orthodox in theory. However, he pointed out that the new method had a wider range of applicability than either of those and that it required less time for computation than the method of fitting constants.

Snedecor and Cox[3] compared the three methods on a series of small samples drawn randomly from a specified population. They stated that differences among the results obtained were trivial with the possible exception that the mean squares for main effects computed by the method of weighted squares of means were usually somewhat smaller than the corresponding values obtained by the other two methods. They concluded that the method of expected numbers can safely be applied to data in which the sub-class numbers do not deviate significantly from proportionality as judged by chi-square. The only case in their experience, they said, in which the results were not in essential agreement with those of the method of fitting constants was one in which the postulation of proportional numbers in the sub-classes of the population was untenable and for which, in fact, the test for agreement of the observed sub-class numbers to proportionality yielded a highly significant value of chi-square.

Recently the author had occasion to compare the methods of expected numbers and of fitting constants in an analysis of sex differences in the growth rate of swine. The data used were exactly the same in both cases; no data were included in the analysis by fitting constants that would have involved sub-classes containing no actual observations. Data from seven non-interbreeding lines of one breed were involved. A separate analysis was made of the data from each line; they

---

* Paper No. 2083, Scientific Journal Series, Minnesota Agricultural Experiment Station.

† Now at North Carolina State College.

[1] George W. Snedecor, "The Method of Expected Numbers for Tables of Multiple Classification with Disproportionate Subclass Numbers," this JOURNAL, 29: 389–393, 1934.

[2] F. Yates, "The Analysis of Multiple Classifications with Unequal Numbers in the Different Classes," this JOURNAL, 29: 51–66, 1934.

[3] George W. Snedecor and Gertrude M. Cox, "Disproportionate Subclass Numbers in Tables of Multiple Classification," *Iowa Experiment Station Research Bulletin 180*, 1935.

were made in a form to give estimates of variance arising from the sources indicated below.

Source of Variation
- Sex
- Years
- Sires within years
- Litters within sires
- Sex × years
- Sex × sires within years
- Sex × litters within sires
- Pigs within sub-classes

The seven analyses were then aggregated in a single analysis in which variances arising from line differences and the interaction of sex and lines were also estimated.

The method of expected numbers was computed as described by Snedecor with the exception that the expected sub-class numbers were based on a postulate of equal frequency of the two sexes instead of on the actual marginal totals for sexes. This reduced the labor of computation since it yielded numbers involving no fractions other than one-half. Tests for departure from the expected numbers so obtained produced no significant values of chi-square. The sex frequencies were very close to equality in six of the seven lines. (Table I.) It should be noted that the above departure from usual procedure does not change the estimation of the sex differences.

TABLE I

THE TOTAL NUMBERS OF EACH SEX IN THE VARIOUS LINES

Line	Number of pigs	
	Male	Female
1	102	100
2	143	114
3	137	127
4	125	125
5	160	159
6	41	37
7	110	122
Total	818	784

The major difference between the results of the two methods and the only one consistent in direction was in the estimates of the sex-litter interaction. The mean squares for this interaction (both methods) are recorded in Table II along with the mean squares for the differences within sub-classes. In every case the interaction mean square is highest when calculated by the method of expected numbers. Considering the single analysis of the data from all lines taken together the results of that method accepted at face value indicate a probability of less than 0.1 that the sex-litter interaction observed was entirely the consequence of random errors. The work on these data was not begun with the intention of comparing analytical procedures but the conclusions that might have been based on this particular result of the method of expected numbers were considered sufficiently important to justify reanalysis by another method. The probability of the amount of interaction present arising from random errors is greater than .05 according to the method of fitting constants.

TABLE II

MEAN SQUARES FOR SEX-LITTER INTERACTION OBTAINED BY DIFFERENT METHODS OF ANALYSIS

Line	Interaction mean square		Within sub-classes mean square	Degrees of freedom	
	Expected numbers	Fitting constants		Interaction	Within sub-classes
1	.0307	.0273	.0136	27	134
2	.0301	.0217	.0241	34	167
3	.0235	.0197	.0162	32	186
4	.0305	.0233	.0132	26	182
5	.0161	.0142	.0183	29	241
6	.0256	.0161	.0289	8	52
7	.0239	.0219	.0249	33	150
1–7	.0257	.0210	.0188	189	1112

Yates stated that the method of fitting constants provides an efficient test of interaction in a two way table. Snedecor and Cox say that it should be looked upon as the standard of comparison. The evidence indicates then that in the case being reported the sex-litter interaction was over estimated by the method of expected numbers.

The sum of squares for interaction in the analysis of a two way table is, symbolically, $\Sigma f(\bar{x}_{ij}-\bar{x}_{r_i}-\bar{x}_{c_j}+\bar{\bar{x}})^2$ where $\bar{x}_{ij}$ is the mean of the subclass in the i-th row and the j-th column, $\bar{x}_{r_i}$ is the mean[4] of the i-th row, $\bar{x}_{c_j}$ the mean[4] of the j-th column, and $\bar{\bar{x}}$ the grand mean.[4] Assuming no true interaction to exist each separate squared deviation multiplied by the appropriate f is an estimate of $d/n\ \sigma^2$ where d is the number of degrees of freedom for the interaction, n the number of squared deviates, and σ^2 the random variance of individual observations. Summing the n values and dividing by d yields an estimate of σ^2. If the value obtained has only a small probability of having arisen by chance alone from the same population as the variance estimated from the variation within sub-classes, reasonable doubt is cast on the assumption of no interaction and it may be concluded that a real interaction exists.

---

[4] As estimated by the method of analysis being used.

The appropriate f for each sub-class mean in the expression given in the preceding paragraph is the sub-class number (k) since the variance of a mean is σ^2/k. The procedure of substituting the expected number (k') for k results in each value of $f(\bar{x}_{i_j}-\bar{x}_{r_j}-\bar{x}_{c_j}+\bar{\bar{x}})^2$ estimating $\frac{k'}{k}\cdot\frac{d}{n}\sigma^2$ instead of $\frac{d}{n}\sigma^2$. The net effect, on the average, is overestimation. The amount to be expected on this basis can be found by averaging k'/k over all the sub-classes. This was done for the data concerned and the results recorded in Table III along with the amount of overestimation indicated by comparison with the results of fitting constants. It will be noted that the agreement is only fair for the separate analyses. There is no apparent method for judging what degree of concordance could reasonably be expected. However, the apparent lack of bias indicated by the remarkably close agreement in the case of the analysis of all the data taken together is favorable evidence for the suggested explanation of the differences in the results of the two methods.

TABLE III

CALCULATED OVERESTIMATION OF SEX-LITTER INTERACTION MEAN SQUARE BY THE METHOD OF EXPECTED NUMBERS AND ACTUAL BY COMPARISON WITH THE RESULTS OF FITTING CONSTANTS

Line	Overestimation	
	Expected	Actual
1	.151	.125
2	.267	.386
3	.265	.195
4	.157	.314
5	.164	.129
6	.299	.584
7	.203	.087
1–7	.210	.222

Following the same reasoning applied in considering the estimation of variance from interaction it appears that the estimates of variance from class differences by the method of expected numbers should also be too large. Assuming the class differences to be purely random this estimate should accordingly approach $\frac{1}{\Sigma(1/k')}\cdot\Sigma\left(\frac{1}{k}\right)\sigma^2$ averaged over all the classes instead of σ^2. Because, in the case at hand, the expected numbers are equal for the two sub-classes of each litter, the overestimation from litter differences should be the same as from the sex-litter interaction. This is in seeming disagreement with the results obtained; in the aggregate of the seven separate analyses the variance estimated from litter differences is .0600 by the method of expected numbers and .0569 by fitting constants. The difference is only about 5 per cent of the latter figure. However, to the extent that the deviations were real, substitution of expected for actual sub-class numbers is immaterial. It is only the random portions of the squared deviations of the class means that vary in average magnitude with the numbers of observations on which the means of the sub-classes of a class are based. The correct expectation then is that the variance estimated from litter differences obtained from the method of expected numbers will exceed that from fitting constants by 21 per cent of the random variance, .0188. This amount is .0039 compared to the actual difference, .0031.

It should be noted that the amount of overestimation of variance from class differences or interaction decreases rather rapidly as the sub-class numbers increase in size so long as they deviate from proportionality only as a result of randomness. The case herein reported is about as extreme as would occur within that limitation.[5] However, there will be in the application of the method to the general run of data a certain frequency, perhaps quite small, of instances like the one herein reported in which this overestimation will, if not recognized, have a bearing on conclusions drawn.

While the existence of real class differences or real interaction reduces the percentage overestimation of the respective mean squares the matter is not critical under those circumstances anyway since the conclusion that class differences or interactions are real is not then in error.

SUMMARY

The actual sub-class numbers in the data concerned did not deviate significantly, as judged by chi-square, from the expected numbers. However, the method of expected numbers by comparison with that of fitting constants overestimated the mean square for the sex-litter interaction by enough to have a bearing on interpretation of the results and the mean squares for litters by a slight amount. The differences in results of the two methods was interpreted to be the consequence of shifts in weighting squared deviations associated with the shift from observed to expected numbers. This led to expected results in reasonable agreement with those observed.

---

[5] Snedecor stated "The change from actual to expected numbers is merely a matter of weighting, usually of minor consequence unless carried to extremes." Certainly this is an extreme case. However, in his next sentence Snedecor implied that adjustments great enough to affect the results would not occur when the chi-square test did not indicate significant deviation of the observed numbers from proportionality.

The overestimation of mean squares reported should not be interpreted as a serious defect of the method of expected numbers. When the condition laid down by Snedecor for the fit of expected numbers to proportionality is met, the maximum overestimation of class and interaction mean squares is not large and the type of data for which a very close approach to this maximum occurs should not be overly frequent. Certainly the advantages cited by Snedecor when contrasted with the defect discussed commend the method for wide usage. What has been presented should actually assist in judging in critical cases whether bias from the source considered is important, thereby permitting increased rather than decreased confidence in the usefulness of the method.

UTILITY OF STATISTICAL METHOD IN AERODYNAMICS

By Herbert G. Smith

As is undoubtedly the case in many a new field of endeavor, it is quite natural, perhaps, that aerodynamics has borrowed several of its physical laws from other sources. Most of these borrowed laws, as well as many of those developed wholly or largely in this field, admittedly are empirical in character in that they have been established in experimental surroundings, generalizing from relatively limited facts. And, just as surveyors readily acknowledge the necessarily approximate nature of *any* measurement, so are most aerodynamic engineers quick to concede that these physical laws are not established firmly beyond question.

Published materials dealing with aerodynamics, particularly the numerous Reports, Technical Memoranda and Technical Notes which have been issued by the National Advisory Committee on Aeronautics, indicate clearly that the wind tunnel has been the most productive source of data from which some of the more recent general physical laws of aerodynamics have tentatively been established, and it is equally clear that applications of the calculus have been among the more frequently used tools for analyzing and developing such data.

Surprising as it may appear to those familiar with statistical methods of analysis, and with the wide applicability of these methods, observation indicates that beyond the occasional use of a few of the more elementary methods for compiling and arranging data, and for obtaining averages, these methods are practically unused in aerodynamic circles, either for establishing or testing laws or for predicting performance. One of statistical method's most powerful devices, correlation, is indicated to have been used little if at all for such purposes. One reason for this lack of use of the excellent analytic tools of statistical methods is the fact that many engineers seem either unaware of their existence and availability or are not familiar with their applicability; another possible reason is that because of the relatively few data (as few as four or five observations, in some instances) from which empirical laws are established, statistical methods are admittedly not always appropriate, being more suitable, of course, for handling and treating masses of data.

Empirical laws and actual-performance data.—Even aerodynamic engineers, themselves, are not always fully satisfied with results obtained by the empirical laws they employ. One capable aerodynamist informed the writer that every aircraft has its own particular characteristics which must be ascertained in flight. Differing somewhat, another engineer stated rather gloatingly that his company has the reputation of building planes conforming more closely to pre-fabrication predictions than is the general case. Thus, there appears ample justification for questioning some of the fundamental aerodynamic laws—particularly in the terms in which most commonly stated—and for suggesting other, supplementary methods for research and analysis.

One example of a fundamental law which may legitimately be questioned in the form usually stated is that established by Daniel Bernoulli (1700–1782), now popularly adapted to airfoils: As velocity of a given volume of a fluid increases, pressures decrease. But aviation engineers with whom the writer has discussed this law invariably hasten to qualify the general, popular version of it to the following: For a flow of a given volume of a fluid, *total* pressure remains constant regardless of velocity; as velocity increases, however, *dynamic* pressures increase, and *static* pressures decrease.

Having been advised by several competent aerodynamic authorities that most published actual-performance data on airplanes are reasonably comparable, representative and reliable, the writer has made use of these generally available (except on recent ships) and relatively numerous data, rather than the more limited and less available wind tunnel data, in making certain analyses by the statistical methods of correlation. (These methods are, of course, quite as suitable for use with laboratory data as with actual-performance data.) The general applicability, as illustrated by the results of two such analyses reported here, may be of interest to statisticians, indicating, as they do, a potentially fertile field; also, should these pages come to the attention of aerodynamic engineers, it is hoped that this discussion may prove thought-provoking, indicating, as it attempts to do, how statistically derived laws may supplement empirically derived laws to advantage.

How lift depends on speed.—The first illustration: Consider the empirical law which, in general terms, holds that lift varies directly as the square of speed. To test this law, actual-performance data on 52 randomly selected and reasonably comparable military monoplanes were studied. Power loading[1] was correlated against speed and against the square of speed; and wing loading[2] was likewise correlated against speed and against the square of speed. The following results were obtained:[3]

Power loading × Speed, $r = -0.643 \pm 0.0548$

For estimating power loading from speed,

$\bar{Y} = 19.53 \text{ lbs.} - (0.0353 \text{ lbs.} \times \text{speed}) \pm 2.02 \text{ lbs.}$

---

[1] Pounds of gross weight per horsepower.

[2] Pounds of gross weight per square foot of supporting area.

[3] The measure of reliability employed here in all instances is the probable error, or 0.6745 of the standard error.

Power loading×Speed squared, $r = -0.604 \pm 0.0595$
For estimating power loading from speed squared,
$\bar{Y} = 14.85 \text{ lbs.} - (0.0000617 \text{ lbs.} \times \text{Speed squared}) \pm 2.11 \text{ lbs.}$
Wing loading×Speed, $r = +0.832 \pm 0.0288$
For estimating wing loading from speed,
$\bar{Y} = 5.52 \text{ lbs.} + (0.0775 \text{ lbs.} \times \text{Speed}) \pm 2.48 \text{ lbs.}$
Wing loading×Speed squared, $r = +0.795 \pm 0.0344$
For estimating wing loading from speed squared,
$\bar{Y} = 15.69 \text{ lbs.} + (0.000138 \text{ lbs.} \times \text{Speed squared}) \pm 2.71 \text{ lbs.}$

Whether lift is considered in terms of power loading or in terms of wing loading, an examination of the r's and their probable errors clearly indicates that, contrary to empirical practice, it is somewhat more logical to consider lift in terms of unaffected speed than in terms of speed squared. In each instance, the $\bar{Y}$ obtained will serve, within reason, as a "law" of how, for the type of airplane considered, etc., lift may be expected to vary with speed or with the square of speed as the case may be.[4]

Horsepower requirements.—As a second illustration of the applicability of correlation methods for testing aerodynamic hypotheses and establishing laws for prediction purposes, consider how horsepower requirements vary with speed. The results of this investigation, hitherto unpublished, are presented briefly.

On the accompanying figure, for purposes of comparison horsepower per gross pound has been tallied respectively against speed, against speed squared and against speed cubed. The $\bar{Y}$-curve in each section may prove useful for predicting how horsepower requirements depend on speed, speed squared or speed cubed, respectively; the $\bar{A}$-curve may prove useful for predicting speed from horsepower; the $\bar{B}$-curve, for predicting speed squared from horsepower; and, the $\bar{C}$-curve for predicting speed cubed from horsepower. The dashed lines outline the probable errors of the prediction curves. The performance data employed are for the same 52 military monoplanes for which the lift analysis (supra) is reported.[5] The results:

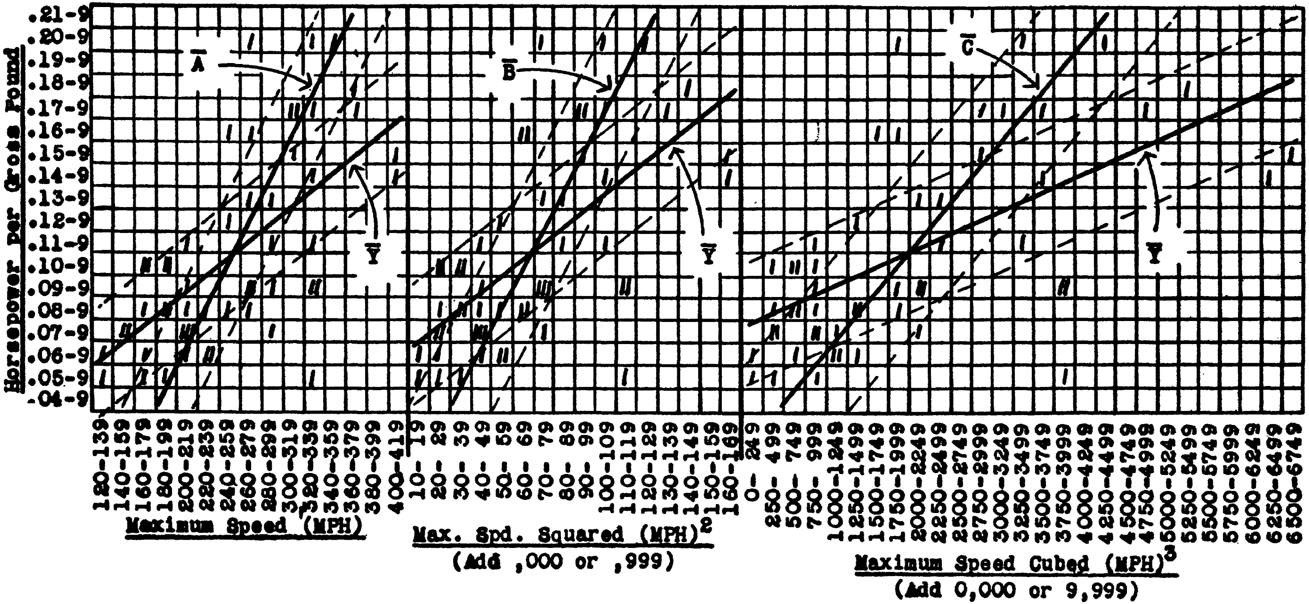
For Horsepower×Speed, $r = +0.650 \pm 0.0541$
$\bar{Y} = (0.009735 + 0.0003975\text{A}) \pm 0.02235$
$\bar{A} = (136.41 + 1062.0\text{Y}) \pm 36.2$
For Horsepower×Speed squared, $r = +0.644 \pm 0.0548$
$\bar{Y} = (0.06008 + 0.000000734\text{B}) \pm 0.02250$
$\bar{B} = (6{,}710.0 + 562{,}000.0\text{Y}) \pm 19{,}650.0$

---

[4] For additional details and comments on the foregoing study, see the writer's "How Lift Varies with Speed," *Aero Digest* for February 1943.

[5] Most data on current aircraft are restricted for military reasons. The data employed here appeared variously in print early in 1942, and may be said to be for ships now largely obsolete.

HOW HORSEPOWER REQUIREMENTS VARY WITH AIRPLANE SPEED



For Horsepower×Speed cubed, $r = +0.619 \pm 0.0575$

$\bar{Y} = (0.07618 + 0.000000001704C) \pm 0.02305$

$\bar{C} = (-4{,}737{,}500.0 + 225{,}000{,}000.0Y) \pm 8{,}410{,}000.0$

Here the correlation coefficients again agree closely, each being moderately high; yet, considered in connection with their delimiting measures, it is clear that for the data analyzed, it is more valid to consider horsepower requirements dependent upon speed unaffected than it is to consider them dependent upon speed squared, or upon speed cubed. In short, it is indicated here that the empirical aerodynamic law, "horsepower varies with the cube of speed," must not be depended upon too greatly.

The following data, also developed, may prove of incidental interest:

Average horsepower, 52 aircraft, 0.1106 ± 0.0294, per gross lb.
Average speed, 254.0 ± 47.7 M.P.H.
Average speed squared, 68,900 ± 25,700
Average speed cubed, 20,150,000 ± 10,700,000

The validity of these findings is further supported by the consistency of indications determined in making a considerable number of other analyses of similar nature.

While the methods of simple, or linear, correlation cannot logically be expected alone always to supply all desired information, the writer's studies lead him to believe that aerodynamic engineers could gain much contributory information were these and other methods of statistical analysis to be made at least supplemental to their presently used analytical procedure. It may be said that when these methods are extended to include partial and multiple correlation techniques some of the objections which might legitimately be raised are obviated. For example, "lift" is not dependent upon speed alone; drag and other factors are influencing factors. Thus, by these extensions of the methods of simple correlation the influences of several factors may be considered —the influence of one or more may be eliminated, or their joint-effect may be weighed. Too, under some instances it is possible that some types of data may be analyzed most satisfactorily by employing the methods of curvilinear correlation.

ON MEASURES OF DISPERSION FOR A FINITE DISTRIBUTION

BY ALBERT O. HIRSCHMAN

IT IS OBVIOUS that the measures of central tendency and of dispersion could be applied, for descriptive purposes, to various section or subgroups of frequency distributions. In particular, the mean and the median often divide distributions into two groups each of which may deserve separate investigation. The dichotomy of language often leads us to speak of "good" and "bad" students, of "rich" and "poor" countries, of "prosperous" and "depressed" industries, and if these expressions cannot be defined by reference to objective standards of scholastic achievement, of national wealth or of industrial prosperity we have to define them by reference to the mean or the median of the relevant statistical series. If we then ask questions about the average good student or the typically depressed industry, we are speaking in terms of the subseries created by the bisection of the entire distribution by mean or median. It is the aim of this paper to show some very simple algebraic relationships between measures for these subseries and the usual measures of dispersions for entire distributions. In the course of this inquiry we shall also discover a simple algebraic expression connecting the mean deviation and standard deviation of the entire distribution.

DEFINITIONS

Let X_1, X_2, $\cdots$, X_n be a series ordered by size of which no item coincides with either the arithmetic mean M or the median Md.[1] Let the number of elements smaller than the mean be i, the number of elements larger than the mean $n-i=s$. If the median does not coincide with any one item in the series n is necessarily an even number. Then, by convention, the median is defined as the average of the $(n/2)$th and the $(n/2+1)$st item. On the other hand, when $s=i=n/2$, the arithmetic mean may be anywhere between the two central items.

We define now

(a) as "partial means"

(1) M_1 and M_2 as the means of the items below and above the median (Md)

(2) M_i and M_s as the means of the items below and above the mean (M)

---

[1] For the contrary case our results would be intrinsically unchanged though algebraically somewhat more complicated.

(b) as "partial standard deviations" (and their squares as "partial variances")
   (1) σ_1 and σ_2 as the standard deviations of the items below and above the median
   (2) σ_i and σ_s as the standard deviations of the items below and above the mean

Between the mean and the partial means we can immediately establish the following relationships:

$$M = \frac{M_1 + M_2}{2} \quad \text{and hence} \quad M - M_1 = M_2 - M \tag{1}$$

$$M = \frac{iM_i + sM_s}{n} \quad \text{and hence} \quad (i + s)M = iM_i + sM_s, \quad \text{and}$$

$$i(M - M_i) = s(M_s - M). \tag{2}$$

EXPRESSION FOR MEAN DEVIATION FROM MEAN AND MEDIAN

We shall now express the mean deviation from the median and from the mean in terms of the "partial" means and the general mean.

(1) Mean deviation from the median (C_{Md})

$$C_{Md} = \frac{\sum_1^n |X - Md|}{n} = \frac{\sum_1^{n/2} (Md - X) + \sum_{n/2+1}^{n} (X - Md)}{n}$$

$$= \frac{n/2Md - \sum_1^{n/2} X + \sum_{n/2+1}^{n} X - n/2Md}{n}$$

so that

$$C_{Md} = \frac{M_2 - M_1}{2} \tag{3a}$$

and according to (1)

$$C_{Md} = M - M_1 = M_2 - M. \tag{3b}$$

(2) Mean deviation from the mean (C_M)

$$C_M = \frac{\sum_1^n |X - M|}{n} = \frac{\sum_1^i (M - X) + \sum_{i+1}^n (X - M)}{n}$$

$$C_M = \frac{iM - \sum_{1}^{i} X + \sum_{i+1}^{n} X - sM}{n}$$

so that

$$C_M = \frac{i(M - M_i) + s(M_s - M)}{n} \tag{4a}$$

and according to (2)

$$C_M = \frac{2i}{n}(M - M_i) = \frac{2s}{n}(M_s - M). \tag{4b}$$

The formula (3a) reminds us immediately of the quartile deviation, which is given by half the distance between the two quartiles Q_3 and Q_1. The quartiles can be interpreted in our terminology as "partial medians" in the same sense as we have called M_1, M_2 and M_i, M_s partial means. The mean deviation from the median and, to a lesser extent, the mean deviation from the mean—note (3b) and (4b)—are thus seen to be refined "ranges"; i.e., distances between certain points lying on the abscissa of a statistical distribution. It is sometimes said in statistical textbooks that an advantage of the quartile deviation over the other measures of dispersion is its simple meaning.[2] This can hardly be maintained with respect to the mean deviations from mean and median, which are shown to be essentially measures of the distances between the general and the partial means. In addition the expressions (3b) and (4b) may be of practical use in the computation of these measures of dispersion especially when the data are ungrouped. Expressions (4a) and (4b) of course reduce to (3a) and (3b) when $i = s = n/2$.

EXPRESSION FOR SKEWNESS

The relationships which we have established also yield a simple measure of skewness. Analogous to the common measure

$$\frac{(Q_3 - Md) - (Md - Q_1)}{(Q_3 - Md) + (Md - Q_1)}$$

relating to the differences between median and quartiles we can indeed form the differences between the median and the partial means M_1 and M_2 and compute the ratio of the difference and of the sum of these differences. This gives

---

[2] See Yule and Kendall, *Introduction to the Theory of Statistics*, London, 1937 (11th edition) p. 149.

$$Sk = \frac{(M_2 - Md) - (Md - M_1)}{(M_2 - Md) + (Md - M_1)}$$

$$= \frac{M_1 + M_2 - 2Md}{M_2 - M_1}$$

which becomes, according to (1) and (3a),

$$Sk = \frac{M - Md}{C_{Md}}. \tag{5}$$

This expression has, as just shown, a meaning analogous to that relating to the quartiles and, like it, varies between 0 and ±1. On the other hand, it resembles the often used measure $\frac{3(M - Md)}{\sigma}$ in that it depends on the position of the mean and therefore on every item in the distribution.

RELATION BETWEEN THE STANDARD AND MEAN DEVIATIONS

In order to show the relationship between the standard deviation and the "partial" quantities defined above we write, according to a well-known property,

$$\sigma_1^2 = \frac{\sum_{1}^{n/2} (X - M)^2}{n/2} - (M - M_1)^2$$

$$\sigma_2^2 = \frac{\sum_{n/2+1}^{n} (X - M)^2}{n/2} - (M_2 - M)^2.$$

Adding and dividing by 2

$$\frac{\sigma_1^2 + \sigma_2^2}{2} = \frac{\sum_{1}^{n/2} (X - M)^2 + \sum_{n/2+1}^{n} (X - M)^2}{n}$$

$$- \frac{(M - M_1)^2 + (M_2 - M)^2}{2}.$$

According to (3b) and to the definition of σ, this can be written

$$\sigma^2 = C_{Md}^2 + \frac{\sigma_1^2 + \sigma_2^2}{2}. \qquad (6)$$

Similarly we derive from

$$\sigma_i^2 = \frac{\sum_1^i (X - M)^2}{i} - (M - M_i)^2$$

$$\sigma_s^2 = \frac{\sum_{i+1}^n (X - M)^2}{s} - (M_s - M)^2$$

the following alternative expression for the variance

$$\sigma^2 = \frac{i(M - M_i)^2 + s(M_s - M)^2}{n} + \frac{i\sigma_i^2 + s\sigma_s^2}{n}$$

$$= \frac{n}{4i} \cdot \frac{4i^2(M - M_i)^2}{n^2} + \frac{n}{4s} \cdot \frac{4s^2(M_s - M)^2}{n^2} + \frac{i\sigma_1^2 + s\sigma_s^2}{n}.$$

According to (4b) this gives

$$\sigma^2 = C_M^2 \cdot \frac{n^2}{4is} + \frac{i\sigma_i^2 + s\sigma_s^2}{n}. \qquad (7)$$

As before, the expression involving the "partial" quantities relating to the median is simpler than that involving the "partial" quantities relating to the mean, and the latter reduces to the former when $i = s = n/2$. Both expressions (6) and (7) show in a novel form the characteristic difference in behavior of mean deviation and standard deviation for different distributions.

The mean deviations from both mean and median depend only upon the distances of the partial means from the general mean. (Quantities which, by analogy to the partial standard deviations, could be defined as partial mean deviations, do *not* enter into the expressions for C_{Md} and C_M.) The standard deviation, on the contrary, is a function of both these differences *and* of the dispersion of the elements around the partial means.[3] It is at once evident from these expressions that the standard deviation can be equal to the mean deviation only in the exceptional case of a symmetrical and ideally U-shaped distribution, in which $i = s$ and in which the partial variances are zero.

---

[3] The coefficient $n^2/4is$ of C_M in (7) shows that, in addition, the standard deviation increases when, with mean deviation from the mean and partial standard deviations constant, the skewness of the distribution increases. In the expression (6) this element is taken care of directly by the quantities σ_1 and σ_2, which can *simultaneously* become 0 only if $i = s$, whereas this is not the case for σ_i and σ_s.

The knowledge of this difference in behavior may help in choosing between the rival measures of dispersion in statistical work *of a purely descriptive character* (i.e., not concerned with sampling problems). Thus in educational statistics, if we speak about dispersion of the marks of a group of students, we may have in mind merely the distance between the average good and the average bad student, or we may want to take account also of the dispersion within good and within bad students. Only in the latter case should the standard deviation be used, whereas in the former case one would have to use the quartile deviation, mean deviation from median or mean deviation from mean, according to the type of average selected.

It is also seen that the knowledge of mean and standard deviation of a distribution tells us something about the partial standard deviations. By writing expressions (6) and (7) in the following form

$$\sigma^2 - C_{Md}^2 = \frac{\sigma_1^2 + \sigma_2^2}{2} \tag{8a}$$

$$\sigma^2 - C_M^2 \cdot \frac{n^2}{4is} = \frac{i\sigma_i^2 + s\sigma_s^2}{n} \tag{8b}$$

we obtain an average of the two partial variances weighted by the frequencies to which each variance relates (since these frequencies are by definition always $n/2$ for both σ_1 and σ_2, a simple average is justified in (8a)).

In the case of a typically symmetrical distribution, not only do the two expressions coincide but we should have in addition $\sigma_1 = \sigma_2$, so that the difference of the squares of standard deviation and of mean deviation would yield directly the partial variance.

RELATION TO GEARY'S TEST OF NORMALITY

As was stated in the beginning, this note is limited to the field of descriptive statistics. In one respect, however, the result reached may be of interest to sampling theory. It is known that for the normal curve we have

$$\frac{\text{Mean deviation from mean}}{\text{Standard deviation}} = \sqrt{\frac{2}{\pi}} \quad 0.79788$$

Now this ratio, which may be denoted by the letter a, has been proposed as a test of normality by R. C. Geary, who has derived its approximate sampling distribution and calculated tables of the mean value, the standard error, and 10, 5 and 1 per cent probability levels of a for samples from a normal population.[4] This test was meant to supplement the usual tests $\sqrt{b_1}=\frac{m_3}{m_2^{3/2}}$ and, in particular, $b_2=\frac{m_4}{m_2^2}$, where m_2, m_3 and m_4 denote the moments taken from the mean.

These two criteria are easily connected with the shape of the frequency distribution, $\sqrt{b_1}$ standing for symmetry and b_2 for kurtosis. The relationship which we have established makes it possible to interpret the a-test in a similarly concrete way. Supposing a symmetrical distribution ($i=s=n/2$), and dividing by σ^2, expression (8b) becomes

$$\frac{C_M^2}{\sigma^2} = a^2 = 1 - \frac{1}{2}\left(\frac{\sigma_i^2}{\sigma^2} + \frac{\sigma_s^2}{\sigma^2}\right)$$

and if symmetry is such that $\sigma_i=\sigma_s$, this reduces to

$$a^2 = 1 - \frac{\sigma_i^2}{\sigma^2}\cdot$$

For symmetrical distributions the ratio of mean deviation to standard deviation is thus negatively related to the partial standard deviations, taken in units of σ. A decrease in a therefore means an increase of σ_i/σ or σ_s/σ, and for unimodal symmetrical distributions this is possible only by a lengthening of the tails of the distribution compensated by other elements drawing closer to the center so as to keep σ constant, i.e., by a development of the distribution in the leptokurtic direction. We are thus able to understand intuitively the high negative correlation between the b_2 criterion standing for kurtosis and Geary's a which E. S. Pearson and Geary have established experimentally and theoretically.

---

[4] R. C. Geary, "The Ratio of the Mean Deviation to the Standard Deviation as a Test of Normality," *Biometrika*, Vol. 27 (1935), pp. 311–332; E. S. Pearson, "A Comparison of b_2 and Mr. Geary's w_n Criterion," *ibid.*, pp. 333–352; R. C. Geary, "Note on the Correlation between b_2 and w_n," *ibid.*, pp. 353–355; R. C. Geary, "Moments of the Ratio of the Mean Deviation to the Standard Deviation for Normal Samples," *ibid.*, Vol. 28 (1936), pp. 295–305; R. C. Geary and E. S. Pearson, *Tests of Normality*, Biometrika Office, London, 1938.

ON SOME CENSUS AIDS TO SAMPLING

BY MORRIS H. HANSEN AND W. EDWARDS DEMING
Bureau of the Census

SOME OF THE AIMS IN SAMPLING DEVELOPMENT

SAMPLING speeds results, reduces costs and demands on manpower, and reduces the burden of reporting. These are factors of prime importance in times of stress, like the present, when situations must often be evaluated and acted upon within a brief space of time. Sampling is playing an indispensable role by providing information for many and various administrative and planning programs under pressure and requirements of accuracy that a few years ago could not have been met. The gains in sampling practice that have proved to be so useful in war time will be in equal demand in time of peace and readjustment.

The purpose of any survey, sample or complete, is to provide factual evidence that presumably will be useful in formulating a course of action in the solution of some problem. There is thus a certain degree of accuracy required in an estimate that is to be obtained from a sample. Consequently, the aim desired in designing a sample is to devise procedures that will

i. Operate within the available budget and limitations of time and manpower;
ii. Operate also within other imposed administrative limitations or restrictions; yet
iii. Produce the maximum amount of information that is possible within the limitations i and ii.

    To meet this requirement, the plans for the survey will call for the use of any available equipment, personnel (organized field or office staff), and other facilities (maps, available statistical data, existing knowledge of the universe, and the like). Even though this requirement be met, the sampling plan is still not satisfactory unless evidence can be produced in the form of mathematics coupled with experience to show that it will

iv. Give results that are reasonably sure to fall within a certain allowable sampling error, which for any survey will be predetermined by the administrative uses that are to be made of the data.

A sample design is judged by the way in which it measures up to these requirements. One design may be better than another in one requirement, and worse in another. A sample design is improved by any device that produces an advantage in meeting one or more of these requirements.

With the aim of improving the efficiency of its surveys, the Census has been conducting a sampling research and planning program for compiling and investigating the distributions of various characteristics of universes to be sampled, evaluating the effectiveness of alternative sampling procedures and designs, developing new theory and adapting existing theory to meet the needs of new requirements, and measuring experimentally not only the sampling variances encountered, but also the time and cost factors involved in alternative methods of collecting data. This research and planning program is supplemented by the preparation of sampling aids in the form of summary cards for small areas, some of which are described in the second half of this paper.

On the basis of recent advances in theory and practice in the Census and elsewhere, and on the basis of accumulated experience in conducting censuses and surveys of various kinds, gains are being made in the specification of sample designs that provide the most information possible under budgetary and other administrative limitations that are always imposed, yet utilizing existing facilities to the fullest degree. It is now possible to lay out satisfactory sampling designs to provide factual data in the solution of many kinds of administrative problems. Further improvement in such designs will be made concurrently with

i. Improved availability of statistical information already on hand. This may take the form of
   a. Punched card summaries of statistics for small areas, for use in stratification and sample selection. An example is the E.D. Summary Card (*vide infra*) prepared from the Census of 1940.
   b. Analytic summaries of the distributions, variances and covariances of certain characteristics of the universe, computed for different sizes and kinds of sampling units, and for diverse characteristics of the universe. Such information is needed in deciding the size and kind of sample to be taken.

ii. Additional theory (particularly concerning systematic sampling designs, type and method of selection of sampling units, errors of response, stratification, etc.).

iii. Additional experience and records regarding field work, interview and travel costs; reactions of the respondents, and the effects of these reactions on the inferences to be drawn from the data collected; comparisons between what was actually done in the field, and what was intended to be done; additional experience in wording questionnaires and training interviewers.

SOME RECENT ADDITIONS TO SAMPLING AIDS

Although the stress of war conditions has produced severe disruptions in population and occupational distributions, data from the Census of 1940 provide a mass of detail that is invaluable in the design of many sampling projects that deal with population or agricultural items. More recent figures on the distribution of certain population characteristics would be desirable, and the estimates of net migration furnished by registrations and rationing programs help to provide these. However, data from the 1940 Census, particularly with regard to the characteristics of small areas as distinguished from characteristics of individuals in the population are extremely useful and effective in designing current sampling projects. In view of the usefulness of the 1940 Census data, we here describe three sets of summary cards for small areas that are valuable as sampling aids. Special tabulations of these cards can be provided at cost.

The E. D. Summary Card.—The E. D. Summary Card[1] is an example of information derived from the 1940 Census, recently made available on punched cards. Through this card, data for small areas are accessible for purposes of stratification and sample selection, and for certain summaries. The card was devised for the purpose of producing useful controls for sampling in various kinds of surveys—housing, population, agriculture, labor force characteristics, marketing, etc. An E.D. Summary Card has been punched for every E.D. outside the "block cities." The characteristics punched on the E.D. Summary Card are enumerated below.[2]

---

[1] The abbreviation E.D. denotes enumeration district. An E.D. is one of the small administrative areas into which the entire country is divided for the purpose of taking a census. Every point of the country lies in one but only one E.D. An E.D. is delineated with the intention that it shall contain a population no greater than one enumerator can cover within the enumeration period. It is further required, however, that no E.D. shall cross the boundary of any incorporated place, city ward, census tract (in a tracted city), township, or other minor civil division, no matter how small. The usual range in the population of an E.D. is between 500 and 1500, though because of the second of the two requirements just noted, many E.Ds. contain populations much smaller than 500—some even as low as one or two persons, or none at all. A small E.D. is usually assigned to an enumerator who is to cover another one nearby, with the consequence that there are more E.Ds. than enumerators. As a matter of fact, in the 1940 Census there were 154,000 E.Ds., but only 110,000 enumerators.

[2] The starred items (*) are ratios that have been calculated for the entire minor civil division of which the particular E.D. forms a part. Every E.D. in a minor civil division shows the same minor civil division ratios. A minor civil division is the smallest political area recognised in the census. Incorporated places are separate minor civil divisions, and outside of incorporated places, townships, election precincts, beats, etc. are used, the nature of the M.C.D. varying from one state to another. Each minor civil division contains one or more E.Ds., but in many instances the M.C.D. and the E.D. are coextensive.

State and county
Enumeration district
Minor civil division
Metropolitan district
Cropland density in the minor civil division*
Population density in the minor civil division*
Urban-rural code
Total population
Native white male population
Per cent native white male
Negro male population
Per cent Negro male
Farm population
Per cent farm
Total number of dwelling units
Number of owner-occupied dwelling units
Per cent owner-occupied
Average rent per dwelling unit (based on rent paid in rented dwelling units, and estimated rental value of owner-occupied dwelling units)
Average farm acreage code for the minor civil division*
Average farm value code for the minor civil division*
Total number of farms

(For explanation of * see footnote 2.)

The Block Summary Card.—It was just stated that the E.D. Summary Card was punched for E.Ds. outside block cities.[3] It was not necessary to punch the E.D. card within block cities, because there existed already the Block Summary Card for each block in these cities. The block is ordinarily a much smaller unit than the E.D. and hence better for most sampling purposes. The block cards correspond to the statistics published in the Supplement to the First Series of Housing. The characteristics punched on the Block Summary Card are enumerated below.

City
Ward
Enumeration district
Block
Number of structures
Total dwelling units
Owner-occupied dwelling units
Tenant-occupied dwelling units
Vacant dwelling units for sale or rent
Vacant dwelling units, other
Dwelling units reporting year built
Dwelling units built 1930–1940
Dwelling units built 1920–1929
Dwelling units built 1900–1919
Dwelling units built 1899 or before
Occupied dwelling units
Dwelling units occupied by nonwhite households
Dwelling units reporting persons per room

---

[3] A "block city" is a city for which statistics by blocks have been issued. This has been done in the Supplement to the First Series Housing Bulletins (16th Census: 1940) for the 191 cities that had 50,000 or more inhabitants in 1930.

Dwelling units with 1.51 or more persons per room
Dwelling units reporting on plumbing equipment and repairs
Dwelling units needing repair or with no private bath
Dwelling units needing repair
Dwelling units with no private bath
Dwelling units reporting on mortgage
Dwelling units mortgaged
Dwelling units reporting rent
Total monthly rent
Average monthly rent per dwelling unit

The County Summary Cards.—A number of cards have been prepared to bring together significant summary information by counties. Additional county information is assembled from time to time on new cards as required in current sampling studies. Most of the data on the cards are already available in published form, but the sources are scattered. The county cards provide data immediately accessible for sampling use.

In addition to the summary cards for counties, a set of cards has been prepared to provide primary sampling units consisting of one, two, or three contiguous counties. The 3067 counties of the country have been consolidated into approximately 2000 primary sampling units. In forming a primary sampling unit, contiguous counties are grouped together so as to give a unit that is as internally heterogeneous as possible. Such consolidations have been introduced for the purpose of providing sampling units more efficient than single counties, and are intended for use when the number of administrative centers is limited and the sampling must be confined to a field of operation around each administrative center. As a matter of fact, the Census made these consolidations in the process of revising the sample used for the Monthly Report on the Labor Force and for many special surveys. Approximately 60 administrative centers are used for this sample.

THE SOCIAL INSURANCE MOVEMENT

By R. Clyde White*
University of Chicago

Social insurance is concerned with some of the risks which are familiar in the field of private casualty insurance. One type of casualty insurance provides indemnity for losses or damages arising out of the legal liability of the insured for loss or damage to the person or property of another. The second type provides benefits in the case of loss or damage to the person or property of the insured. Most social insurance belongs to the second type, but workmen's compensation should probably be classified under the first because the employer is legally liable to pay compensation to his injured employee. Hence, social insurance may be regarded as public casualty insurance, authorized by law and usually administered in whole or in part by government and designed to maintain some portion of the income of the insured. A social insurance law defines a risk, prescribes a formula for the determination of the amount of cash or other benefits and provides for the collection of contributions to pay the costs of benefits. A contribution tax is generally levied on the employer or on the employer and the employee. Frequently the law requires an appropriation out of general governmental revenues as a subsidy to the system. Coverage is always defined in reference to an employed person, sometimes including self-employed persons, and benefits are frequently payable to the dependents of insured persons under specified conditions.

In large measure the enactment of social insurance laws in a nation has followed the growth of large scale industrial and commercial enterprise. Historically the roots of social insurance reach back many centuries, but these early experiments were conducted by small mutual aid groups. About the end of the eighteenth century some municipal governments in Europe set up very limited forms of social insurance, but the laws were permissive and usually not very successful. In 1786, John Acland proposed a general plan of social insurance in England as a means of relieving the taxpayer of his relief burdens. The project was discussed in Parliament, but no legislative action was taken. Tom Paine advocated something of the sort in this country. It was not until 1883, however, that a national legislative body adopted a social insurance law which was applicable by compulsion to large numbers of workers. This was the famous sickness insurance law sponsored by Bismarck

---

* The writer wishes to acknowledge the painstaking work of Miss Eloise Whitney and Mrs. Lois Sentman, two of his graduate students in the University of Chicago. They spent many hours examining the sources and tabulating the material of Table I in this article.

in Germany. Agricultural and domestic service workers are frequently covered by social insurance laws, but in general the concern of the national legislative bodies has been with industrial and commercial employees.

Social insurance may be voluntary or compulsory. If it is voluntary, it offers certain benefits to the contributor or to the potential beneficiary which he may take or leave. A law which prescribes a compulsory scheme defines the employments covered and names the beneficiaries by industrial or occupational class. Those involved have no choice. Under most compulsory laws there are likely to be certain classes of persons who may elect coverage, but they represent a small percentage of the population insured. Voluntary social insurance was for several decades characteristic of Belgium, France, The Netherlands, Scandinavia and Switzerland, but during the last fifteen years these countries have been changing to compulsory coverage. Compulsory coverage gives greater assurance that all persons subject to a risk will be continuously covered and, therefore, protected, and experience has shown that benefits are likely to be more nearly adequate to assure minimum subsistence under this type of law.

The aim of this study has been to chart the course of social insurance legislation and to determine what laws are now in operation in all parts of the world. In order to simplify this task the phrase, "social insurance law," has been rather narrowly defined. As used in this paper, a social insurance law is one which was enacted by a national legislative body for the protection of a class or classes of workers which may include government employees—such as civil service workers, teachers, policemen, etc.—but must include others. Pension legislation for war veterans and their dependents and war risk insurance which many governments have had and still have in force have been excluded, because they relate to the special condition of war and not to the ordinary peace-time risks. A departure has to be made from the strict concept of national legislation in the case of the United States, because the courts have held that the states have the power to enact protective legislation such as social insurance. For example, the state of Washington adopted the first general workmen's compensation act in this country in 1911 (the Federal Government enacted a law in 1908 but it covered only federal employees) and Wisconsin enacted the first unemployment compensation law in 1932. No attempt is made in this paper to show changes in the original law, unless it was entirely repealed, which has occurred only once, and that was in the case of the Russian unemployment insurance law.

Obviously public assistance, as it is known in English speaking

TABLE I

SOCIAL INSURANCE LAWS BY COUNTRY OR DEPENDENCY, RISK COVERED AND DATE OF ORIGINAL ENACTMENT*

Country or dependency	Risk covered and date of original enactment						
	Accident	Sickness	Invalidity	Unemployment	Old age	Survivors	Maternity
Africa							
Italian Colonies	1931†	—‡	—	—	—	—	—
Mozambique	1931†	—	—	—	—	—	—
Rhodesia	1931†	—	—	—	—	—	—
South Africa	1931†	—	—	—	—	—	—
West Africa	1940	—	—	—	—	—	—
Asia							
British India	1923	—	—	—	—	—	1941§
China	1929	—	—	—	—	—	—
Iraq	1937	—	—	—	—	—	—
Japan	1911	1922	—	—	—	—	—
Palestine	1927	—	—	—	—	—	—
Philippines	1927	—	—	—	—	—	—
Australasia							
Australia	1900	1938	1908	—	—	1938	1938
New Zealand	1908	1938	1938	1938	1938	1938	1938
Europe							
Austria	1887	1888	1925	1920	1925	1925	1888
Belgium	1903	1903	1924	1933	1924	1937	—
Bulgaria	1908	1924	1937	1925	1924	1924	1924
Czechoslovakia	1888	1888	1924	1921	1924	1924	1888
Denmark	1898	1898	—	1907¶	—	—	1937
Esthonia	1923	1911	—	—	—	—	1911
Finland	1895	—	1937	1917	1939	—	—
France	1898	1928	1928	1905	1938	1928	1928
Germany	1884	1883	1889	1927	1889	1911	1883
Great Britain	1897	1911	1911	1911	1925	1925	1911
Greece	1901	1934	1932	—	1932	1932	1934
Hungary	1907	1927	1928	—	1928	1928	1927
Irish Free State	1906	1911	1911	1911	—	1935	1911
Italy	1898	1928	1923	1919	1923	1923	1935**
Latvia	1927	1911	—	—	—	—	1911
Lithuania	1938	1925	—	—	—	—	1925
Luxemburg	1902	1901	1911	—	1911	1911	1901
Netherlands	1901	1913	1913	1916	1913	1913	1913
Norway	1894	1909	—	1906¶	—	—	1909
Poland	1920	1920	1927	1924	1927	1927	1920
Portugal	1913	1919	1919	—	1919	—	1919
Roumania	1912	1912	1912	—	1912	—	1912
Russia	1903	1922	1922	—	1922	1922	1922
Spain	1900	—	—	1919	1939	—	—
Sweden	1901	1935	1913	1934	1913	—	1935
Switzerland	1912	1911	—	1924	—	—	1911
Yugoslavia	1922	1922	1922	—	1909	1922	1922
Islands, Pacific							
Fiji Islands	1940	—	—	—	—	—	—
Hawaii	1915††	—	—	1937††	—	—	—
New Caledonia	1931‡‡	—	—	—	—	—	—
North America							
Canada	1902§§	—	—	1940	—	—	—
Costa Rica	1925	1941	—	1941	1941	1941	1941
Cuba	1916	—	1940	1940	1940	1940	1939

TABLE I (Continued)

Country or dependency	Risk covered and date of original enactment						
	Accident	Sickness	Invalidity	Unemployment	Old age	Survivors	Maternity
Guatemala	1924	—	—	—	—	—	1926
Mexico	1906	—	—	—	—	—	—
Newfoundland	1908	—	—	—	—	—	—
Nicaragua	1930	—	—	—	—	—	—
Panama	1942	1942	1942	—	1942	—	1942
Puerto Rico	1916¶¶	—	—	—	—	—	—
United States***	1911	—	1935	1932	1935	1939	—
South America							
Argentina	1923	—	—	—	1923	—	1936
Bolivia	1924	1939	—	—	1939	—	—
Brazil	1919	1940	1940	—	1940	1940	1940
Chile	1924	1924	1924	—	1924	—	1924
Columbia	1915	1939	—	1939	—	—	1939
Dominican Rep	1932	—	—	—	—	—	—
Ecuador	1928	1935	1935	—	1935	1937	—
Paraguay	1927	—	—	—	1941	—	—
Peru	1911	1941	1941	—	1941	1941	1941
Uruguay	1914	—	1919	1939	1919	—	—
Venezuela	1928	1940	1940	1940	1940	—	1940

* Page citations and dates of publications are not indicated because of the large number of such references, but the data for this table were obtained from the following sources: *Legislative Series*, International Labour Organisation, Geneva and Montreal, 1923 to the present; *International Labour Review*, International Labour Organisation, Geneva and Montreal; *Monthly Labor Review*, U. S. Department of Labor; and Barbara N. Armstrong, *Insuring the Essentials*, The Macmillan Company, New York, 1932.

† Probably some minor legislation earlier.

‡ Indicates absence of this kind of coverage.

§ Covers only women working in mines.

¶ Ghent system only. Belgium and Denmark adopted the Ghent system at early dates also but changed to the more common system on the dates given.

** Includes benefits at marriage and birth.

†† These were territorial laws and were the first such laws for this dependency.

‡‡ Law adopted some time before this date, but exact date could not be established.

§§ First provincial workmen's compensation act.

¶¶ First territorial law.

*** The federal law of 1908 covered only federal employees for workmen's compensation. The date given is the Washington State Act which was the first workmen's compensation law to cover employees other than governmental. In 1942 Rhode Island redefined unemployment to include unemployment due to illness and now pays compensation in such cases, but this is not defined here as sickness insurance. The invalidity insurance covers only railroad employees under very restricted conditions. The Wisconsin Unemployment Compensation Act is shown as the first United States unemployment insurance law.

countries, has been excluded from consideration here. It is not in any ordinary sense insurance. In public assistance the risk is only vaguely defined; the liability of a public fund for public assistance payments is not established, because officials have wide discretion; funds for the payment of public assistance are appropriated out of general revenues and vary according to political attitudes from one legislative session to another; and the potential recipient must submit to a means test. By contrast for the purposes of a social insurance law the risk is carefully defined, the liability of the insurance fund is stated in terms of a benefit formula, funds are derived from earmarked taxes or in other predetermined ways, and the beneficiary does not submit to a means test.

The clearly defined risks which were found in this study are accident (including occupational disease), invalidity of long duration, maternity, old age, sickness, survivors (i.e., generally only widowhood and orphanage) and unemployment. At the beginning of the present war all of these risks were covered in Austria, Bulgaria, Czechoslovakia, France, Germany, Great Britain, the Netherlands, Poland and New Zealand. Table I shows the years in which social insurance laws in force at the end of November, 1942, were first adopted by national legislative bodies. Exceptions in the case of the United States will be indicated. A "national social insurance law" may apply to certain classes of the entire population, or it may permit a political subdivision to enact social insurance legislation and cover certain defined classes. For example, Titles III and IX of the United States Social Security Act did not create an unemployment insurance system in this country, but they carried incentives to the states to adopt such measures. Obviously that was national legislation in the social insurance field, but the date of the Wisconsin Act which preceded the Social Security Act is shown as the first American unemployment insurance law. Dates for laws applying to certain dependencies, such as Hawaii and Rhodesia, are given separately just to indicate the extension of social insurance measures to outlying parts of the world.

Several observations may be made concerning Table I. First, all of the countries and dependencies have accident insurance (workmen's compensation in English-speaking countries) but no other form of social insurance is found in all of them. Second, of the 253 laws counted and listed 149 are found in European countries. That is 58.9 per cent of the total. This is perhaps not unexpected, since social insurance was a European invention. Third, only 18 of the 52 laws adopted between 1930 and 1939 were in Europe. Fourth, from 1940 to the end of 1942 new laws numbering 34 were adopted, and all of them were outside of Europe. Within 12 years the non-European world adopted 68 new social insurance laws. It appears to be clear that this European invention has now been accepted in principle and to a large degree in fact as a useful social device in most of the world. Many population groups in the countries which have various types of social insurance are not covered, but historically the tendency has been to broaden coverage once a type of social insurance was adopted, and it is reasonable to expect that extension of coverage will continue. At present only two countries approach one hundred per cent coverage of all population groups, and they are Russia and New Zealand. The only large populated areas of the world which have not yet been affected by the social insurance movement are some of the large Pacific islands, southeast Asia, southwest Asia, Turkey and the central mass of Africa. These areas probably do not contain more than 10 per cent of the population of the world, and because of their lack of industrialization social insurance is not well adapted to the economy of these areas.

TABLE II

NUMBER OF NATIONS HAVING SOCIAL INSURANCE LAWS BY RISK COVERED

Risk	Number of nations
Accident	64
Invalidity	31
Maternity	37
Old age	34
Sickness	37
Survivors	24
Unemployment	26

The prevalence of various forms of social insurance is shown in Table II. The risk of industrial accident has been most widely recognized as a hazard which should be covered by insurance, and many countries have only this type of social insurance. Sickness and maternity insurance stand next in frequency of adoption.

The growth of social insurance in time is shown in Table III which shows by quinquennia the dates of origin of particular kinds of social insurance in the various countries. There was a steady increase in the number of new laws adopted in the several countries up to the outbreak of the first world war, but during the period of the war new enactments dropped sharply. The quinquennium immediately following the war was marked by the largest number of new laws of any similar period. The early years of the depression showed a drop in new legislation, but, when recovery began in the middle thirties, the movement regained its momentum.

Nothing in the foregoing tables indicates the extent of coverage in the different countries of a specific form of social insurance, and nothing is said about the adequacy of the benefit provisions. There is great diversity with respect to both groups covered and the adequacy of benefits. The aim of this study was to determine the countries which had accepted the principle of social insurance as a method of achieving

TABLE III
NUMBER OF SOCIAL INSURANCE LAWS ADOPTED, BY QUINQUENNIA

Quinquennium	Number of laws
Total	253
1880–84	3
1885–89	9
1890–94	1
1895–99	5
1900–04	12
1905–09	13
1910–14	35
1915–19	16
1920–24	41
1925–29	31
1930–34	15
1935–39	38
1940–42	34

social security and the kinds of risk which had been recognized as insurable.

The widespread attention, which the recent report on the British social services by Sir William Beveridge received, is undoubtedly due in large part to the general understanding of the insurance principle in all countries and to the common recognition that this principle is applicable to other risks and to other groups in the population than those currently covered. The fact that this report concerns itself almost entirely with social insurance and is uncomplicated by more elaborate proposals for reorganizing the national economy may account for the cordiality of its reception in this country. This contrasts sharply with that accorded the recent report of our National Resources Planning Board, dealing with the same problems. The report of the Planning Board lacks the simplicity and clarity of the Beveridge report, and this lack was undoubtedly due in large measure to the complex program which the Planning Board outlined. The Beveridge report has given new impetus to the social insurance movement.

HORACE SECRIST, 1881–1943

THE SCIENCES of statistics and economics lost a scholar of high rank in the death of Professor Horace Secrist. As a research student, as an author, and as a teacher, he always maintained a rigorous scientific attitude toward his work. His main research activities lay in the realm of statistics, but his interests were not confined to technique, as he was always seeking for the hidden meaning, or relations, that lay back of the factual data. He firmly believed that truth must be discovered from an analysis of factual data and he had little patience with "armchair" economics. If a proposition did not lend itself to statistical verification, he had little use for it.

Secrist was widely read in the literature of his fields of interest. He was acquainted with the writings of the leading scholars both at home and abroad, and he was constantly searching the literature of scientific research for proven methods of establishing the truth. Regardless of the character of the statistical problem that was occupying his attention, whether it be some governmental question that required an immediate answer, or a problem that was perplexing a group of business men, he never lost sight of his scientific interest in the data. Often after the immediate problem had been disposed of, an article would appear from his pen which disclosed that the data had been subjected to further analysis for the purpose of discovering what contribution they might make to our understanding of economic relations.

Secrist was always a rigorous teacher. He had set high standards of attainment for himself and he expected similar ones from his students. Many a student who complained, while in college, of Secrist as a hard task-master, soon learned upon entering a business career that the discipline he had received in Secrist's classes was his best asset. Graduate students always benefited greatly by work with him. He opened their eyes to the rigors of scientific method and developed within them a wholesome respect for scientific caution.

He was author of a large number of books and monographs dealing with statistical problems, but these problems were always treated with an eye for their social and economic implications. In addition to his own scholarly contributions, he was active for many years in the work of the Social Science Research Council, having served as its secretary for several years. He frequently turned his statistical knowledge and skill to the solution of some governmental problem. In my judgment, he developed the most effective method of determining "bulk-line cost" as a means of adjusting price to output in war times. If this method had not been kicked out of governmental councils in our present war emergency, we would now be hearing less about shortages, "roll-backs" and subsidies.

For years Secrist suffered from a physical disability that finally caused his death. Nevertheless during all this time he maintained his scholarly activity and produced some of his best work. Throughout his career he lived the life of a true scientist and is an example to be emulated by others who wish to contribute to a better understanding of our economic and social life.

F. S. DEIBLER

COMMITTEE ON NOMINATIONS

President Goldenweiser has appointed the following members as the Committee on Nominations for 1943: F. Leslie Hayford, formerly Economist of the General Motors Corporation, New York City, *Chairman;* Professor F. L. Carmichael, School of Commerce and Bureau of Business and Social Research, University of Denver, Denver, Colorado; and Professor Douglas E. Scates, Duke University, College Station, Durham, North Carolina. The report of the Committee on Nominations will be published in the November BULLETIN.

BOOK REVIEWS

GLENN E. MCLAUGHLIN

Review Editor

War without Inflation, by George Katona. New York: Columbia University Press. 1942. x, 213 pp. $2.50.

A more informative title for this book might have been "The Art of Anti-Inflationary Propaganda." For its main thesis is that with the economic weapons available for fighting inflation, certain psychological conditions must be fulfilled if the fight is to be successful. To be more explicit: The experience of every belligerent country has shown that it is practically impossible to close the "gap" by fiscal measures alone. Therefore, it becomes necessary that if inflation is to be avoided there must be a substantial increase in the rate of voluntary saving. For, if the public were determined to spend its pre-war proportion of its disposable income, direct controls would be powerless to prevent inflation. For an adequate rate of voluntary saving to exist, it is necessary that the public should have confidence in the Government's anti-inflationary program. To the extent that confidence is lacking, the rate of spending must increase.

There is, of course, nothing novel in this general thesis, but Governmental technique does not always indicate an awareness of it. Indeed, as the author points out, the actions of public spirited people with the best intentions often contribute to the wrong psychological attitude being produced. For the greater the eminence of the administrator warning against inflation, the more reason the public has for believing in its imminence. Also, the requirements of sensational journalism tend to produce the wrong impression. If the President announces that measures are to be taken to avoid a potential increase in the cost of living of 25 per cent, the resulting headline will probably be "25 per cent increase in living costs feared." The way to avoid all this completely is as obvious as it is sinister; but a substantial contribution could be made if the press would sacrifice its headlines, which it could without sacrificing its freedom.

The book is full of interesting suggestions but not conclusions as to the best psychological approach. For instance, in organizing a voluntary saving campaign, is it more effective to dramatize the need for sacrifice or to say that it is more sensible to save now in order to secure an automobile or a radio in 1947? The author makes the important point that a tax program should be suitable to impress the public that the Government is determined to fight inflation. With the tax prospects of an election year, this may be the main contribution that the 1944 taxes will make.

While this book hardly deserves to be required reading for administrators, it should stand high on the "recommended" list.

A. SMITHIES

Washington, D.C.

Outlay and Income in the United States, 1921–1938, by Harold Barger. New York: National Bureau of Economic Research. 1942. xxvii, 391 pp. $2.50.

It is a truism that the value of the national product during a given period may be determined by summating either the values of the services rendered in production, or the values of the goods and services constituting the final product. The standard estimates of the national product of the United States are estimates of the income received by the agents employed in its production. Mr. Barger is not the pioneer in estimating independently the aggregate outlay for the final product. Lough and Warburton, to mention no others, have preceded him. But Barger presents the first essentially independent estimate of outlay which covers a series of consecutive years. It is therefore possible for the first time to compare income and outlay estimates as series, and national income literature is the richer. The purpose of the volume, the author states, is to contribute to materials for the study of the business cycle. Comparison of outlay and income estimates, and analysis of relations between components, may, however, be as useful to other students of the national income as to cycle analysts.

In addition to annual and quarterly estimates in current dollars of income and outlay for the period 1921–1938, the volume presents a stimulating discussion of concepts, useful running commentary on methodology, and speculation both qualitative and statistical concerning sources and margins of error in the estimates. Throughout, the study evidences the work of a fertile mind.

The annual data for income, and with one bold exception the income concepts, are Kuznets'. Kuznets' interest is basically in the social product as such; Barger's is avowedly in variation over the cycle. To the reviewer, it seems that only failure to adjust the data to the change in purpose can account for retaining Kuznets' treatment of depreciation and depletion, inventories, and capital gains and losses. Since in each case the alternative figures are presented in subsidiary tables, the point is perhaps not important. But Barger is certainly not correct in asserting that, given consistency, the concepts adopted are immaterial. On the controversial matter of the product of government, Mr. Barger "out-Herods Herod" by assuming that all government product covered by taxes *of any kind* is intermediate product, and deficit expenditures alone measure the value of the final product of government. As a measure of total product in the system, this treatment seems indefensible, even as a choice among available alternatives. But considering the purpose of the volume, there may be a better defense for Barger's treatment than any defense which he presents. Any study of causal forces in cyclical fluctuations must in one way or another relate movements in output or activity totals to changes in the level of savings and investment. But the total to be so related is the total of output net rather than gross of tax-covered public product. For, to a first approximation, the propensity to save of government is zero, and public tax-supported activity is therefore not a part of the total which is uniquely related to the current level and rate of change of savings. Public and private activity are not alternatives. Rather, public tax-supported activity is superimposed upon the level of private activity.

Not the least attractive and stimulating feature of the volume is the frankness and clarity with which Mr. Barger discusses the shortcomings of the data. The statistical analysis of probable margins of error is also valuable though it is a little startling to see the outlay and income data for a given year treated as though they were a random sample of two drawn from an infinite population of estimates of the national product.

The statistical analysis might well have been carried further, as Richard Stone has pointed out in a review article.[1] As Stone shows, the residual discrepancies between the outlay and income estimates, after a constant difference has been subtracted, are significantly correlated with change from the preceding year in physical volume of inventory, with profits from inventory revaluation, and (perhaps not significantly) with time. Here are promising clues for further investigation.

Stone points out, too, that Barger is hasty in concluding that the two series cannot be brought together into one. They cannot be reconciled in the usual sense, it is true. But a weighted average may be secured which presumably would be more valid than either of the series averaged. It "would be possible on the lines suggested by [Stone] and others to assess measures of precision for the component series which by suitable combination would yield approximate weights for the two measures of the national product, and also weights for distributing the adjustment in each series over its component items."[2] The single series so derived might be highly useful in further statistical analysis of cyclical behavior.

EVERETT E. HAGEN

Washington, D. C.

International Trade and the National Income Multiplier, by Fritz Machlup. Philadelphia: The Blakiston Company. 1943. xvi, 237 pp. $3.50.

In recent years it has become common to meet in economic discussions the formidable terminology of the *multiplier*, to the acute discomfort of the uninitiated. Professor Machlup has attempted to repair the deficiency of elementary expositions of this subject, and if any conscientious reader completes this book without a fair notion of what is involved in the multiplier concept, it will not be because of any lack of pedagogical zeal on the part of the author. The subject is developed in painstaking, and at times almost painful, detail. Although the discussion concentrates upon the relationships between countries, the theory of the domestic multiplier is developed by implication, so that the book will also appeal to those who have no special interest in international trade as such.

By avoiding the use of all but the simplest algebraic symbolism, the author has succeeded in compressing into a book of some 200 pages what could otherwise be presented in a single mathematical article of some complexity and limited appeal. Numerous arithmetical model sequences are presented under a variety of assumptions, involving variations in numerical coefficients and in the initial disturbances assumed. These models will tax the energy of the reader, but if the essential principles of the first one are understood, then so should be the other thirty or so, including the "super-duper" on page 100.

---

[1] "Two Studies on Income and Expenditure in the United States," *Economic Journal*, April, 1943 (Vol. LIII), pp. 60–75.

[2] *Loc. cit.*, p. 74.

There are nine chapters and two mathematical appendices. Chapters I, II, VIII, and X are not very technical in nature. In these, Professor Machlup discusses the background of the multiplier models, assumptions, qualifications, policy implications, etc. The reviewer was not satisfied at every point with the author's analysis: e.g. the monetary explanation of the workings of the multiplier; the discussion of the Robertsonian period definitions of saving and investment; the time length of the interval between successive "rounds"; the belief that multiplier models must assume continuing streams of autonomous change; the brief discussion of capital flight; the largely terminological difficulties conjured up in the discussion as to whether the trade balance causes foreign investment or vice versa; the statistical and conceptual problems in empirical verification of the multiplier hypothesis. Nevertheless, there is so much that is excellent in these pages, and the subject matter is so important, that no student of the cycle can neglect to read and reread them.

The remaining technical chapters constitute the heart of the book, and although more complex, they nevertheless make simpler reading. With many variations, three main cases are studied: (1) the effect of a shift of demand from domestic to foreign products, upon the home country and the foreign country; (2) the effect of increased investment in one country upon both countries; and (3) the effect of simultaneous (opposite) changes of investment in both countries upon each country. The first is worked out in greater detail, for more than two countries, with and without savings, for many combinations of coefficients, in terms of the transient response of the system as well as the limiting final state, with algebraic analysis and elementary difference equation theory. Most of this is embodied in the masterly Chapter V and its two appendices. The reviewer could have wished, however, that the author had explicity used the fact that cases (3) and (1) are really identical except for bracketing of terms. Actually, case (2) is the most important of all, not only empirically but also because the others can be derived from it by simple and natural superposition. By concentrating upon it, the length of the book could have been halved leaving room for further discussion of price and interest changes, imperialism, and other policy problems. Perhaps, too, the book would have been more useful if the author had expounded in greater detail the contributions of Metzler, Salant, and Laursen. However, as it is, the book is an important contribution which deserves to be in the library of all modern economists.

PAUL A. SAMUELSON

Massachusetts Institute of Technology

Economic Equilibrium, Employment and Natural Resources, by L. R. Nienstaedt. Bloomington, Indiana: The Principia Press, Inc. 1942. xvi, 412 pp. $4.00.

This book will appear strange to many economists in that much of the content deviates from that found in traditional analyses of economic problems. An attempt is made to develop a reasonably complete statement of macrocosmic balance in terms of factors involved in physical production. Balance is conceived as an adjustment between production and consumption and derived by relationship to agricultural output in a Physiocratic framework. The relationship to agricultural output is conceived to vary with the stage of industrialization. The system is developed on three major relationships: one—the agricultural harvest to other economic production; two—human and machine capacity to production; three—physical factors to the flow of money values in producing national income. Nienstaedt denies the validity of market-price adjustment. He sees the weakness of microcosmic balance in an economy which varies as widely as ours has in recent years.

A series of relationships is developed and checked against statistical information relating to the United States' economy, principally for the period from 1880 to 1929. The checking of this system against the statistical data of the United States is too involved to be summarized here. One critical measurement may be noted because of its significance in Nienstaedt's denial of the validity of market-price adjustment. The ratio of currency in circulation to national income is found to be fairly stable within given periods. Stated in the simplest terms, national income is derived from production on the one hand and related to the flow of money values by a constant proportion to currency in circulation on the other, so that market-price adjustment is held to have little determining influence. To achieve this strange conclusion it is necessary to ignore the influence of bank deposits.

Nienstaedt holds that equilibrium developed in terms of physical production makes possible the complete integration of the business cycle with the system of equilibrium itself. He thinks of the business cycle as being made up of random causes of the Slutsky effect, time lags as suggested by Rosenstein-Rodan, and the flow of velocities at individual production stages as suggested by Amorosa.

One of the conclusions reached is that balance under present conditions can be achieved only where a considerable part of agricultural production enters the market in free exchange with all other products. This leads to a suggestion of "Industrial Village Estates" as a method of stabilizing employment. Of major importance in his system is the balance of capacities. Therefore, another method of stabilization suggested is the registration of technical productive capacities for which certificates may be issued with the requirement that such certificates be obtained before connecting productive machinery.

The book evidences a good deal of work and, although the reviewer is unable to accept most of the material in its present form, he suggests that the study is worthy of the attention of those interested in the development of ideas of macrocosmic balance.

ELMER C. BRATT

Lehigh University

The Shifting and Incidence of Taxation, by Otto von Mering. Philadelphia: The Blakiston Company, 1942. xiii, 262 pp. $3.25.

This book presents a closely reasoned analysis of shifting, incidence, and effects of taxation, "largely built upon [the author's] . . . earlier work, *Steuerüberwälzung*" (1928). For most of the problems, its approach is traditional in the use of partial equilibrium supply-and-demand mechanics, but modern in the application of recently developed geometrical techniques, especially Joan Robinson's. Some of the proofs in Part I, "The General Theory of Tax Shifting," and the proofs in the appendix are developed by mathematical analysis.

Much of the contents will probably be new to the general student of public finance, for Professor von Mering covers an impressively wide range of problems, utilizing analyses gathered from journals and specialized works both here and abroad. The material is presented not as a history of doctrine but as a fairly well integrated series of problems, the author selecting in each case the source that he evidently considers the best for that topic. The twenty pages on tax shifting under a monopoly, for instance, draw successively on Böhm-Bawerk, Menger, Conigliani, Joan Robinson, Cournot, Edgeworth, Gilbert, Fagan and Jastram, Seligman, H. G. Brown, Dalton, Garver, Stamp, A. L. Meyers, Wicksell, and Hotelling; and include an original treatment by von Mering himself of the difference between the effects of a per unit tax and a gross receipts tax set at a rate to equal the per unit tax (at the price prevailing before tax).

The attempt to cover so many difficult problems in so compressed a space makes real understanding of some of the analysis difficult, in some places probably impossible. Part of the difficulty can be traced to a lack of explicitness in the assumptions underlying the analysis; part, to the use of specialized concepts (like "the investing period," pp. 122-24) that the author expects the reader to be familiar with already or to assimiliate by a study of the original sources. Aside from these two points, however, the book is remarkable for its clarity and directness of expression; the author wastes no words.

Part II, "The Shifting of Particular Taxes," is not quite up to the level of Part I. The reasoning seems to lapse occasionally from its generally high level of dependability (or perhaps it is merely that the assumptions are less explicit), especially in the tendency to emphasize the extensive margin but not the intensive (e.g. example on p. 166; and cf. Edgeworth, Volume II, p. 83). It takes a good deal of caution and trouble to get safely through parts like pp. 184-88, where the "supply of labor" evidently refers to changes in the number of the working force, in the "elastic" case, and to changes in the percentage of the working force that is employed, in the "inelastic" case (hence the meaning of the conclusion that the workers bear the tax is not clear), while the inelastic "demand for labor" seems to run in terms of money wages, in contrast to the real-wage analysis in the preceding cases. And in weighing the conclusions reached by the author, it must be recalled that even in the analysis of widespread taxes like those on all retail sales or on wages in general, the fact that some disposition has to be made of the tax revenue is usually excluded from the scope of the analysis.

For those who have a special interest in the economic aspects of public finance the book will give a substantial reward, if it is absorbed slowly and carefully.

Carl S. Shoup

Columbia University

Federal Tax Course, by George T. Altman. Chicago: Commerce Clearing House. 1943. Loose-leaf. $10.00.

Although intended more for study than as a reference work, this course has been so designed as to serve both purposes, and to provide readily the answers to everyday questions. The entire federal tax structure together with its functioning is adequately covered, but the book rightly lays the greater stress upon the Federal income tax. Accompanying the course, which consists of some 1400 or more pages in a well-indexed and tabbed loose-leaf binder, the publishers furnish a special book containing the income, estate, and gift tax provisions of the Internal Revenue Code as amended to date, and also a separate book of filled-in tax return forms.

The course analyzes not only the statutes, but more significantly, Treasury regulations and Supreme Court decisions as they modify the basic law. The course is divided into four main parts: (1) explanatory text; (2) problems and solutions; (3) specimen income-tax returns; and (4) official texts of the income tax and excess profits tax regulations. The closely packed 11-page glossary of tax terms, the topical index, and the table of cases cited and discussed in the explanatory text round out the day-to-day usefulness of this book.

The explanatory text traces the Federal revenue system back to the early colonies. It shows how the tariff as a source of revenue was first supplemented, and later supplanted by, the internal revenue, which now leans heavily upon income and profits taxes, but with no inconsiderable support from excises, gift and estate taxes, employment taxes, and other miscellaneous internal revenues, many of which started as "temporary" taxes and finally became an integral part of our tax structure.

The text then analyzes the Internal Revenue Code and "other sources of the law," meaning tax laws not incorporated in the Code, Treasury Department regulations and rulings, decisions of the Tax Court of the United States (formerly known as the United States Board of Tax Appeals), and decisions of other courts. Also included here is a short section on the power of the states to levy income taxes.

The main part of the book concerns itself with the federal income tax as it applies to individuals and corporations. Some of the topics dealt with, to mention but a few, are: classes of taxpayers, taxable incomes, rates and exemptions, business expenses, depreciation, depletion, amortization, interest, taxes, medical expenses, and so forth. The remainder of the explanatory text covers excess profits and capital stock taxes, excises, sales and processing taxes, taxes on facilities and admissions, social security taxes, and finally estate and gift taxes.

The succeeding section of the course presents problems and questions illustrating the salient features of the tax laws. These problems are classified in the same manner as the explanatory text, and are appropriately keyed in with it. Specimen filled-in returns for individuals, partnerships, corporations, estates and trusts, and excess profits, plus sample worksheets and explanatory instructions, increase the practical value of this volume and its accompanying materials.

The broad coverage of the course and its ready application extend its usefulness beyond the student to the lawyer, accountant, and other professional worker, as well as to those in government and business.

JAMES D. PARIS

Great Neck, New York

Inter-American Statistical Yearbook, 1942. Edited by Raul C. Migone, with the assistance of Marcelo Aberastury, Emilio Fuente, and Jorge E. Iturraspe, under the auspices of the Comision Argentina De Altos Estudios Internaciales. New York: The Macmillan Company; Buenos Aires: El Ateneo; Rio de Janeiro: Freitas Bastas & Cia., 1942. 1066 pp. $10.00.

The *Inter-American Statistical Yearbook, 1942*, represents the second edition of the publication, the first having appeared in 1940. As in the case of the first edition it was prepared under the Argentine Commission of High International Studies and edited by Dr. Raul C. Migone.

A distinctive feature of the book is that the data in the tables are presented in the four principal languages spoken in the Western Hemisphere, in order of their importance: Spanish, English, Portuguese, and French. This makes the information readily available for reference by any of the four lingual groups, and should, in time, serve as a stimulus toward an understanding and appreciation of the other languages with which the majority of readers are not so familiar as they are with their own.

The major arrangement of the book is by subjects, rather than by countries. This affords a ready comparison of the importance of each area in terms of a given subject. The principal subject headings are as follows: Population, production, industries, commerce (foreign), social questions, currencies, banks, and investments, public finance, public education, public health, international cooperation, and sources. Although most of these subject headings are self-explanatory, there are two which deserve special mention—social questions and international cooperation.

The data presented under social questions are especially valuable in making comparisons of living standards, and of reaching an understanding of how each country utilizes its natural and human resources. These data include information regarding the gainfully employed population, wages and hours, unemployment, strikes, cost of living, retail prices, autosufficiency of foodstuffs, average per capita consumption of the principal foodstuffs, and distribution of expenditures for the more important groups of items in the cost of living.

The section on international cooperation presents in tabular form a complete record of the formal relationships of each country with others, under the following headings: Participation in international or continental organizations, agreements and conventions concluded under the auspices of the League of Nations, Pan-American treaties and conventions, international labor conventions, and bilateral trade agreements. It is unusual to find information of this type arranged in a form so convenient for reference.

A section in the preface is devoted to conversion coefficients for weights and measures used in the different countries, and includes a tabulation of the approximate exchange values of the various currencies. This feature adds much to the value of the book in that the reader does not require additional references to interpret quantities and values.

MYRON E. ANDREWS

Washington, D. C.

World Minerals and World Peace, by C. K. Leith, J. W. Furness, and Cleona Lewis. Washington: The Brookings Institution. 1942. xii, 253 pp. $2.50.

This study was suggested by Dr. C. K. Leith, geologist of the University of Wisconsin, whose interest in the relation of minerals to peace dates from World War I, when he was mineral adviser to the War Industries Board and to the American Commission to Negotiate Peace. Dr. Cleona Lewis of the Brookings Institution staff, who collaborated with the other authors, Dr. Leith and J. W. Furness (formerly chief of the Economics and Statistics Branch of the Bureau of Mines) has written several volumes on international economic relations. The result of this joint effort is a volume that merits the attention of all those interested in raw materials as a factor in world peace.

A substantial background of factual information on trends in mineral production and trade is presented in Part I of the volume. Part II discusses the recent trends in controls through monopolies, cartels, nationalization, and other political and economic policies. The third and final part delves into a necessarily less objective consideration of future access to raw materials and mineral control in relation to peace.

The year 1938 is chosen for most of the recent international comparisons of mineral output. Although that year was the last full calendar year before the outbreak of the major conflict in Europe, the selection is open to criticism on the ground that the sharp recession of industrial activity in the United States that year makes a comparison with other countries less favorable than it would have been for a more nearly normal year, such as 1936 or 1937. However, data presented in Chapter I and Appendix A for certain minerals over a period of years enable an evaluation of the effect of this difference in cyclical position.

Among the more interesting conclusions are that the so-called "have not" nations really did not possess much of the world's mineral production, and efforts to develop domestic supplies met with little success. The proportion of total mineral output controlled by the Axis powers is placed at 11 per cent in 1939 and, as a result of conquest, at 25 per cent at the end of 1942. It is pointed out that the "have nots" were not denied access to minerals. On the contrary, they benefited from keen competition for the sale of minerals in world markets.

The final chapter, written by Dr. Leith, develops the thesis that the use of minerals should be restricted to peaceful pursuits. It is conceded that those who elect to prevent the would-be aggressor must possess most of the minerals. The reviewer confesses that he is left uncertain as to how the follow-up control would be accomplished and still retain the freedom acclaimed by the Atlantic Charter.

Perhaps if the nations in control of major mineral resources had helped the German democracy prosper as readily as they made economic concessions to the might of the Nazis almost from the start, there would not have developed the occasion for the bogus claims that raw materials could not be obtained at the very time that they were being hoarded hand-over-fist in preparation for war.

WILBERT G. FRITZ

War Production Board

Principles of Marketing, by Fred E. Clark and Carrie Patton Clark. New York: The Macmillan Company. 1942. xx, 828 pp. $4.50.

This is the latest—and perhaps the last—general textbook on marketing to be published during the war. It is not that popular interest has waned in the field of distribution but that the exigency of war has shifted emphasis to subjects of industrial production and away from marketing topics. After the war, there will be disproportionment in trained marketing men as there was in trained production men when we entered the war. If experience at the close of the last war follows this war's end, there again will be "boom days" in formal classroom instruction at collegiate schools of business administration.

It is fortunate, then, that Dr. Clark has revised his text on marketing. This third edition appears twenty years after his first edition; it aptly summarizes events of the years in between and touches on war-time economy in marketing only incidentally. Since war-time controls are not likely to be exercised in a peace economy, the experiences of current marketing operations may be left to others to view in a proper historical perspective.

Any student of marketing welcomes Dr. Clark's annotated text. There are footnotes on nearly every page and a bibliography at the close of each chapter. To the casual reader this presentation may deter his speed in reading, but the references serve well to correlate developments in a rapidly changing field which Dr. Clark has ably documented in the period of peace between wars.

To those who are acquainted with the early text, it should be added that the original character of the book has not been changed; the form of treatment is functional. Each chapter has been thoroughly revised with additional material, charts and illustrations. In particular the chapters on retailing have been expanded with special reference to voluntary and cooperative chains and to supermarkets.

The chapters on manufacturers' problems include new material on the marketing of business goods and selective selling. The treatment on manufacturer-middleman relations has been clarified. One chapter has been devoted to the nature of the consumer movement and the various agencies and activities connected with it. Controls and legislation affecting marketing, now so extensive and far-reaching, have been expanded into two chapters. Dr. Clark's text always has been well-oriented on agricultural marketing subjects. The book is divided into twenty-eight chapters together with an appendix on marketing reports which are suggested for classroom problems.

HOWARD T. HOVDE

Wharton School of Finance and Commerce
University of Pennsylvania

The Road We Are Traveling, by Stuart Chase. New York: The Twentieth Century Fund. 1942. 106 pp. $1.00.

In the preface, Stuart Chase writes:

> The United States may come out of this war the strongest nation on earth. To be worthy of that strength, we should take the lead in plans for permanent co-operation among our allies. . . . Our task will begin, even if it does not end, at home.

And then he gives us the first of six exploratory reports on post-war problems under the title *The Road We Are Traveling: 1914–1942*. It is likely that when all six reports have been completed, the Twentieth Century Fund will publish them in a single volume called *When the War Ends*. For just as the course of the war will do much to shape the peace, so intelligent thinking about the peace may affect the course of the war. History, as has been said before, is a seamless process, part idea, part event, with the two interacting at all times.

Since this is the introductory volume to Mr. Chase's projected series, its material is fairly familiar. The first half of the book is a sprightly recapitulation of the events of 1914–1942, a period in which governments everywhere discovered the conveniences of deficit financing and the necessity of giving public employment or subsidies to those who could no longer find work in private industry. One may have misgivings about the whole 1914–1942 pattern, but personal prejudice does not banish the realities which Mr. Chase describes.

The second part of Mr. Chase's book consists of an inventory of basic trends. These are divided into causes and effects, the causes being technological advance, the flattening population curve and the closing frontier. The effects he mentions are interdependence, unemployment, the decline of investment opportunity, shrinkage of the free market, high-pressure selling, mechanized warfare, government in business, and a trend toward national self-sufficiency.

Mr. Chase offers pretty clear going in most of the book but I keep wondering whether his cause-and-effect sequence is entirely adequate. Why should we assume, for example, that technological advance must lead to a decline of investment opportunity? And why should we think that flattening population curves and closed frontiers are here to stay? If government action can prevent the stiffening of prices and the limited use of patents, technical innovations should actually compound the opportunities for investment.

Then too, this war will have a tremendous impact on our attitude toward family sizes; in the future no sovereign nation will dare to let itself be maneuvered into the position of France, where the old began to outnumber the young. With the teeming Orient only a step away, Australia will certainly not be satisfied with a population of 7,000,000. Russia will seek millions of new people to exploit Siberia, which probably will remain beyond effective bombing range from Western Europe for years to come.

Actually, this war should result in a dozen new frontiers. The road to Alaska, if properly exploited as a substitute for such palliatives as WPA and the British dole, should be the equivalent of a minor Louisiana Purchase. For the North Country, as Stefansson persists in pointing out, can support people who are willing to work within its limitations. Our own Pacific Northwest with its unused water power is a potential frontier vast enough to absorb every complaining Okie.

Mr. Chase errs, I think, in assuming that State interference in the economic process must point to more government investment rather than less. We have always had State interference and State subsidies: the Homestead

Act was a gigantic land subsidy, and the building of the American road system was a free gift to that rugged individualist, Mr. Henry Ford. Why can't government interference and subsidization of the future be of a type that will work to free the individual enterpriser as they have in the past? The Farm Security Administration points the way in a modest fashion.

Too often our thinking lags behind our opportunities, and our courage cannot even match our thinking. We monkey around with trying to maintain the status quo in cotton and wheat; but if we rewarded farmers for diversifying their crops instead of giving them money to add to cotton and wheat carry-overs, we might discover shortly that freedom from the necessity of farm subsidies is just around the corner. Has any farm family yet had its fill of tomato juice? At any rate, it is nonsense to suppose that wheat should eternally bring a dollar when every year sees money shaved from the cost of industrial products. Instead of giving the farmer money for taking fields out of use, why not try to give him a cheaper tractor, a more efficient pump?

Inasmuch as he has five more volumes to write, Mr. Chase probably will canvass all possibilities. He may indicate later that he does not consider a vast increase in government investment to be the only post-war alternative before us. Government investment may be one "way-out," but, as Hilaire Belloc says, when the "determining number" of voters depends on the State for livelihood, freedom usually disappears. When political power grows too strong at the expense of social power, democracy usually dies.

FORREST H. KIRKPATRICK

Bethany College

Harvard Cooperative Society, Past and Present, by N. S. B. Gras. Cambridge: Harvard University Press. 1942. xi, 191 pp. $1.50.

Professor Gras has written a fascinating little book detailing the history of a cooperative enterprise in the Harvard environment. In sprightly style, he tells how this organization developed in sixty years a membership equity of $800,000, annual sales of $1,400,000, a total membership of 14,000, and a current dividend rate on members' purchases of 12 per cent. The account is sympathetic, yet well balanced, and does not suffer from lapses into uncritical eulogy.

The significance of the volume goes far beyond the confines of the Cambridge chain of enterprises. As Professor Gras suggests, it is a study of "Twenty years of cooperation and forty years of success" in following the best traditions of private business. But it is also a case study in the encroachment of private business ideology on a pattern of consumers cooperation—an encroachment which has similarly affected thousands of other cooperative enterprises, ranging from mutual insurance companies, through cooperative banks to farmers' telephone companies. In a capitalistic milieu, the democratic idealism of cooperative pioneers has frequently faded from sight.

Compromises and adjustments have been made with surrounding institutions. The result has often been the development of an organization which is neither flesh nor fowl. It is no longer cooperation—perhaps it might be termed a mutual company.

Put briefly, what happened to the Harvard cooperative was this: the Society was launched by students on a fully democratic basis to reduce the cost of living in the Harvard area. It early faced problems of apathy, bad management, and boycotts. For a period, Professor Ames and Professor Taussig, and Harvard itself, gave significant assistance. Substantial growth ensued during the eighteen nineties, accompanied by increasing problems of accounting and control. An apathetic membership was faced with the question: "Should the original democratic way of life prevail or a new aristocratic oligarchy be adopted?"

The oligarchy won and in 1902 the Society super-imposed upon the original organization a complex scheme of voting stockholders in a newly-formed corporation. These voting stockholders, as trustees, were to elect the board of directors, although a sufficient number of membership votes might override the voting stockholders.

Today, the re-assumption of control by the "ticket-holders" is "possible but not practical." The Society is not affiliated with the Cooperative League of the U.S.A. Its reform enthusiasm is dead. Its efficient manager "keeps abreast of the time, partly by attending the meetings of business men who are wrestling with problems similar to his own." He has secured considerable assistance from the Harvard School of Business Administration and the National Retail Dry Goods Association in improving budgetary practice. "The history of the society since 1903 . . . points to success in spite of, or because of a lack of sympathy with, the cooperative movement as generally propagated. One notable example of this has been the lack of any emphasis upon education in cooperation. The Harvard emphasis has been upon private business, notably in its School of Business."

Mr. Gras seems content to place the Harvard Cooperative Society alongside the Ford Motor Company and R. H. Macy and Company as illustrations of financial-industrial capitalism, possessed of managerial efficiency and free from restraining banking influence. He is a bit fearful of stagnation in the organization, of over-emphasis on profitable lines, of advancing costs, and of the failure to enlist the enthusiastic support and participation by students and teachers at Harvard and Tech. He is not pleased by the complicated constitution which "is a model in obscurity." Yet, on balance, he finds the Society "a vital going organization" with "a life, a standing, a reputation, and a function to perform"—the result of conscientious effort of the many distinguished men who have served it well over its history.

To the reviewer, this appraisal appears too favorable. A deeper significance appears in the fact that more than 8,000 members are completely disenfranchised and that but a dozen of the 7,000 full-fledged members appear for ticket-holder meetings. The ten stockholding trustees have the controlling power and have no thought of propaganda or education in cooperation. The Society has achieved acknowledged business success; it has not contributed an object lesson in democracy. Perhaps the latter would be too much to expect in the existing setting.

COLSTON WARNE

Amherst College

Jewish Population Studies. Edited by Sophia M. Robison. New York: Conference on Jewish Relations, 1943. xvi, 189 pp. $3.50.

Since the United States Census does not inquire into the religious affiliation of individuals, and the Census of Religious Bodies also does not comprise an actual enumeration of the numbers of any sect but a compilation from secondary sources, the Jewish Population Studies, initiated by the Conference on Jewish Relations, are a new and important source of information. The studies were planned "on the principle that the communal work of Jewish organizations should be based on facts ascertained by scientific procedure." Demographic analyses of 10 communities (by different authors) are included in the present volume: Trenton (Sophia M. Robison), Passaic (same author), Buffalo (Uriah Z. Engelman), Norwich and New London (Bessie B. Wessel), Pittsburgh (Maurice Taylor), Detroit (Henry J. Meyer), Chicago (A. J. Jaffe), Minneapolis (Sophia M. Robison), San Francisco (Samuel Moment). However, as the editor points out in an introductory chapter on methods applied to gather data on the Jewish population, the results obtained in these cities are likewise based more or less on estimates or sampling procedures "because the arguments not to include the question of religious affiliation in the census blank outweigh those in its favor."

Most of the uncertainties in such studies arise from the difficulty of defining the designation "Jew." In the Chicago study a Jew is one whose death certificate indicates that he was buried by an undertaking firm whose business it is to conduct Jewish burials. In the Detroit study a Jew is one who has a Jewish name. In the other reports Jews are defined as persons born of Jewish parents or of mixed marriages, who are members of Jewish communal and religious organizations, and as those who when interviewed were willing to be identified as Jewish by "race" or religion. Other more methodical difficulties come in by using the estimated birth and death rates or the absence ratio of school children on the high Jewish holiday, Day of Atonement ("Yom Kippur method"), for estimating the total Jewish population on various assumptions. The Buffalo survey only approached a complete census by a house-to-house canvass in those districts known to be densely populated by Jews; in other districts, however, and other places a sample canvass had to be employed.

Despite the differences and deficiencies in method, the results obtained are noteworthy as a good start toward compiling comparable characteristics for the Jewish population in the United States as a whole. The results, for instance, show that the pattern of age distribution, sex ratio, size of family, length of residence of the foreign-born was essentially the same in all the communities studied. Numerous tables give detailed information about these items. When compared with the total white population, the Jewish group in each city showed a smaller ratio of children of elementary school age, a more even distribution among the sexes and a higher ratio of naturalized citizens among the foreign-born.

The most striking characteristic in all ten communities lies in the occupational distribution of the Jews, especially in their concentration in trade and the professions, well-known from findings in other countries. In no city, states the editor in her summarizing conclusions, was the ratio of gainfully occupied Jews in trade less than 43 per cent (Passaic), whereas in Pittsburgh it was as much as 60 per cent. In no one of the cities were there less than 8.5 per cent of all gainfully occupied Jews engaged in professional careers. In two of the cities, Pittsburgh and Norwich, Jews in the professions accounted for more than 9 per cent, and in Trenton and Passaic this ratio went up to 12.3 per cent. Especially high within the professional group was the concentration of Jewish doctors and lawyers. For instance, in Trenton with a Jewish population of 6 per cent, the Jewish physicians and surgeons comprised two-thirds of the 122 persons practicing medicine, while among the 190 lawyers and judges the Jews constituted 40.5 per cent. These results are very similar to those obtained in the last complete census of Germany (1933) regarding the vocational distribution of the Jewish minority and reveal, indeed, a serious problem for a minority group. The careful studies in these American cities although based on estimating and sampling methods deserve the full attention of all circles concerned until a complete enumeration and evaluation by census methods for all religious sects make the results still more representative.

GEORG WOLFF

National Institute of Health

An Introduction to Managerial Business Statistics, by Harry Pelle Hartkemeier. New York: Thomas Y. Crowell Company. 1943. 207 pp. $1.75.

This book reflects the author's extensive experience and firm conviction that the text for the first course in business statistics "should contain thorough and clear statements of a few topics" and should not attempt to cover the whole field in one volume.

There are six chapters in the book: The Introduction outlines briefly what is meant by quantitative, qualitative, frequency-distribution, and time series analysis, index numbers, and correlation and it relates them to problems of business management. Chapter 2 on "Quantitative Analysis" and Chapter 3 on "The Theory Underlying Tests of Significance or Reliability" apply the theory of small samples and reliability to such practical problems as choosing which brand of tires or paint to buy. Chapter 4 treats the subject of "Computing Procedure and Machines." Chapter 5 on "Qualitative Analysis" applies the tests of reliability to such problems as the percentages of customers preferring specific brands, obtained by sampling. Chapter 6 on "Production Control" applies the sampling process to the measurement of the variability of the physical characteristics of products made by a machine or a factory and describes how to set up control charts which show management when these physical characteristics vary beyond the limits of tolerance.

Because these chapters are written so that the subject matter may be understood easily by a student who has not had any college mathematics, the use and derivation of statistical formulae are kept at a minimum. At the end of each chapter many useful work tables and charts are to be found as well as sets of practical business problems.

The pages, 8½ by 11 inches in size, are punched for ring binders. These details have been arranged in order that the book may be taken apart easily so that individual chapters, tables, or instructions may be inserted in a binder along with other student notes or work sheets.

Unlike many other introductory texts, it does not discuss in detail such topics as the planning of statistical investigations, collection of data from various sources by different methods, tabulation, ratios, charting, index numbers or the different measures of central tendency, dispersion and correlation. These methods are also useful in solving business problems. For example, the median and the mode—as well as other statistical measures—have their applications in solving business management problems. Apparently, it is the author's intention to treat index numbers and other topics in a series of separate volumes of which this is the first. The author's criticism of advertising, page 26, seems to be irrelevant, unjudicial, and based on an extremely small sample.

Within the limited range of the subject matter of this book, the author has succeeded in explaining the practical applications of some phases of statistical theory which are usually very difficult for the layman to understand and apply. It should fit well into the curriculum of a business school which endeavors to develop the management point of view. Business executives who possess a good educational background should read this book because it indicates how these methods may be applied profitably to the solution of many of their management problems.

DONALD R. G. COWAN

PUBLICATIONS RECEIVED*

Backman, Jules. *Rationing and Price Control in Great Britain.* Washington: The Brookings Institution, Pamphlet No. 50. 1943. iii, 68 pp. 50 cents.

Berry, Thomas Senior. *Western Prices Before 1861.* Cambridge, Massachusetts: Harvard University Press. 1943. xxi, 645 pp. $5.00.

Butterbaugh, Grant Illion. *The Measurement of Business Activity in the Puget Sound Area.* Chicago: The University of Chicago Press, Studies in Business Administration, Vol. XIII, Number 2. 1943. v, 72 pp. $1.00.

Chlepner, B. S. *Belgian Banking and Banking Theory.* Washington: The Brookings Institution. 1943. x, 224 pp. $2.50.

Crawford, Mary M. *Student Folkways and Spending at Indiana University, 1940–1941.* New York: Columbia University Press. 1943. 271 pp. $3.50.

Dempsey, Bernard W. *Interest and Usury.* Washington: American Council on Public Affairs. 1943. xii, 233 pp. $3.00.

Dublin, Louis I. *A Family of Thirty Million.* New York: Metropolitan Life Insurance Company. 1943. xvi, 496 pp.

Fisher, Ronald A., and Yates, Frank. *Statistical Tables for Biological, Agricultural and Medical Research.* London: Oliver and Boyd Ltd. 1943. viii, 98 pp. 13 shillings 6 pence

Freeman, Frank N., and Wenger, M. A. *The Chicago Mental Growth Battery.* Chicago: The University of Chicago Press. 1943. v, 58 pp. $1.00.

Hartkemeier, Harry Pelle. *An Introduction to Managerial Business Statistics.* New York: Thomas Y. Crowell Company. 1943. xiv, 207 pp. $1.75.

Heaney, Norman Stewart. *Public Trusteeship.* Baltimore: The Johns Hopkins Press. The Johns Hopkins University Studies in Historical and Political Science, Series LX, Number 4. 1942. xiv, 130 pp. $1.50.

Hoyt, Elizabeth E. *Freedom from Want: A World Goal.* New York: Public Affairs Committee, Inc., Public Affairs Pamphlets No. 80. 1943. 31 pp. 10 cents.

Kuczynski, R. R. *The New Population Statistics.* New York: The Macmillan Company. National Institute of Economic and Social Research, Occasional Papers, I. 31 pp. 50 cents.

Labor Research Association. *Labor and the War,* Labor Fact Book 6. New York: International Publishers Company, Inc. 208 pp. $2.00.

Leith, C. K., Furness, J. W., and Lewis, Cleona. *World Minerals and World Peace.* Washington: The Brookings Institution. 1943. xii, 253 pp. $2.50.

Lyon, Stanley. *Some Observations on Births in Dublin in the Years 1941 and 1942.* Dublin: Cahill & Co., Ltd. 1943. 24 pp.

Machlup, Fritz. *International Trade and the National Income Multiplier.* Philadelphia: The Blakiston Company. 1943. xvi, 237 pp. $3.50.

Mangus, A. R., and McNamara, Robert L. *Levels of Living and Population Movements in Rural Areas of Ohio, 1930–1940.* Wooster, Ohio: Ohio Agricultural Experiment Station, Bulletin 639. 1943. 61 pp.

Mayo, Bernard, Editor. *Annual Report of the American Historical Association, 1936,* Vol. III, "Instructions to the British Ministers to the United

---

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. Some items of minor interest to statisticians have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications which are to be found in the *Revue de l'Institut International de Statistique, Journal of the Royal Statistical Society, American Economic Review, Population Index, Transactions of the Actuarial Society of America, The Record of the American Institute of Actuaries,* and *Sankhyā, the Indian Journal of Statistics.*—Editor.

States, 1791–1812." Washington: United States Government Printing Office. 1941. xvi, 403 pp. $1.00.

Migone, Raul C., Editor, assisted by Aberastury, Marcelo, Fuente, Emilio, and Iturraspe, Jorge E. under the auspices of the Comision Argentina De Altos Estudios Internacionales. *Inter-American Statistical Yearbook, 1942.* New York: The Macmillan Company; Buenos Aires: El Ateneo; Rio de Janeiro: Freitas Bastos & Cia. 1943. 1067 pp. $10.00.

Motherwell, Hiram. *Rebuilding Europe—After Victory.* New York: Public Affairs Committee, Inc., Public Affairs Pamphlets, No. 81. 1943. 32 pp. 10 cents.

Moulton, Harold G. *The New Philosophy of Public Debt.* Washington: The Brookings Institution. 1943. 93 pp. $1.00.

National Bureau of Economic Research. *Cost Behavior and Price Policy.* New York: National Bureau of Economic Research, Committee on Price Determination. 1943. xix, 356 pp. $3.00.

National Bureau of Economic Research. *Income Size Distributions, Part II, Studies in Income and Wealth,* Vol. V. New York: National Bureau of Economic Research, Conference on Research in Income and Wealth. 1943. xvii, 1681 pp.

Reeve, Joseph E. *Monetary Reform Movements—A Survey of Recent Plans and Panaceas.* Washington: American Council on Public Affairs. 1943. xiv, 404 pp. $3.25.

Shoup, Carl, Friedman, Milton, and Mack, Ruth P. *Taxing to Prevent Inflation.* New York: Columbia University Press. 1943. viii, 236 pp. $2.75.

Special Libraries Association. *War Subject Headings for Information Files,* Second Edition. New York: Special Libraries Association. 1943. 69 pp. $2.00.

Sturmthal, Adolf. *The Tragedy of European Labor, 1918–1939.* New York: Columbia University Press. 1943. xii, 389 pp. $3.50.

Tax Institute. *Wartime Problems of State and Local Finance.* Philadelphia: Tax Institute, University of Pennsylvania. 1943. ix, 267 pp. $2.50.

Taylor, Henry C., and Taylor, Anne Dewees. *World Trade in Agricultural Products.* New York: The Macmillan Company. 1943. xviii, 286 pp. $3.50.

Tomasevich, Jozo. *International Agreements on Conservation of Marine Resources.* Stanford University: Food Research Institute. 1943. xi, 297 pp. $3.00.

Truesdell, Leon E. *The Canadian Born in the United States.* New Haven: Yale University Press. 1943. xvi, 263 pp. $3.00.

The Twentieth Century Fund. *Postwar planning in the United States, An Organization Directory.* New York: The Twentieth Century Fund. 1943. xvi, 101 pp. $1.00.

Walker, Helen M. *Elementary Statistical Methods.* New York: Henry Holt and Company, Inc. 1943. xxv, 368 pp. $2.75.

Wilks, S. S. *Mathematical Statistics.* Princeton: Princeton University Press. 1943. xi, 284 pp. $3.75.

Willis, J. Brooke. *The Functions of the Commercial Banking System.* New York: King's Crown Press. 1943. viii, 225 pp. $3.00.

Government Publications

Chicago Plan Commission. *Industrial and Commercial Background for Planning Chicago.* Chicago: The Chicago Plan Commission. 1942. 66 pp.

Department of Commerce, Bureau of the Census. *Service Establishments, Places of Amusement, Hotels, Tourist Courts and Tourist Camps, 1939,* prepared under the supervision of Fred A. Gosnell. Sixteenth Census of the United States: 1940 Census of Business, Vol. II. Washington: U. S. Government Printing Office. 1942. xvi, 637 pp. $2.00.

Department of Commerce, Bureau of the Census. *Construction: 1939,* prepared under the supervision of Fred A. Gosnell, and John Albright. Sixteenth Census of the United

States: 1940, Census of Business: 1939, Vol. IV. Washington: U. S. Government Printing Office. 1943. viii, 397 pp. $1.50.

Department of Labor, Division of Labor Standards. *Controlling Absenteeism, A Record of War Plant Experience.* Special Bulletin No. 12. Washington: U. S. Government Printing Office. 1943. 57 pp. 10 cents.

International Labour Office. *The International Standardisation of Labour Statistics.* Montreal: International Labour Office, Studies and Reports, Series N (Statistics) No. 25, Revision of No. 19. 1943. vii, 169 pp. $1.00.

League of Nations. *Relief Deliveries and Relief Loans, 1919–1923.* Princeton: Princeton University Press. 1943. 62 pp.

League of Nations. *The Transition from War to Peace Economy.* Princeton: Princeton University Press: Economic, Financial and Transit Department of the League of Nations. II. A. 3. 1943. 118 pp. $1.00.

National Resources Planning Board. *Human Conservation, The Story of Our Wasted Resources.* Washington: U. S. Government Printing Office. 1943. 126 pp. 20 cents.

Social Security Board. *Measurement of Variations in State Economic and Fiscal Capacity,* by Paul Studenski. Washington: Social Security Board, Bureau of Research and Statistics. 1943. 76 pp.

Social Security Board. *Statistics of Family Composition in Selected Areas of the United States, 1934–36,* Vol. II, The Urban Sample (83 Cities in 18 States). Washington: Social Security Board, Bureau of Research and Statistics, Bureau Memorandum No. 45. 1943. xxxvi, 620 pp.

Union of South Africa. *Annual Accounts, 1941–42 together with Financial Statements Relating to Hospital Boards, Municipalities, Village Management Boards, Local Boards and Divisional Councils for the Year Ended 31st December, 1941, with the Reports of the Provincial Auditor thereon.* Pretoria: Government Printer of the Union of South Africa. 1943. 236 pp.

Imperial
New Delhi

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

DECEMBER, 1943 · $1.50 PER COPY · $6.00 PER ANNUM · VOL. 38 · NO. 224

CONTENTS

ARTICLES

An Appraisal of the U. S. Bureau of Labor Statistics Cost of Living Index
A SPECIAL COMMITTEE OF THE AMERICAN STATISTICAL ASSOCIATION 387

Housing Scales for Rural Pennsylvania . . . HOWARD R. COTTAM 406

Coordinating Production for War . . . . . . SHELBY CULLOM DAVIS 417

Statistics for Wage Stabilization . ROBERT J. MYERS AND HARRY OBER 425

The Distribution of Deposits and Currency in the United States, 1929–1939 . . . . SOLOMON SHAPIRO 438

Variations in Expenditures where Families of Wage Earners and Clerical Workers are Classified by Economic Level . . . . ROY C. CAVE 445

Fitting Polynomial Trends to Seasonal Data by the Method of Least Squares . . . . HOWARD L. JONES 453

A Method for Determining the Significance of a Shortage . . . . LILA F. KNUDSEN 466

Paul T. Cherington, 1876–1943 . . . . . . WILLARD L. THORP 471

BOOK REVIEWS

BARGER, HAROLD, AND LANDSBERG, HANS H. *American Agriculture, 1899–1939: A Study of Output, Employment and Productivity.* Mordecai Ezekiel 484

BERRY, THOMAS SENIOR. *Western Prices before 1861, A Study of the Cincinnati Market.* Ernest W. Williams, Jr. 485

BOWLEY, A. L. (Editor). *Studies in the National Income, 1924–1938.* Milton Gilbert 473

BUTTERBAUGH, GRANT I. *The Measurement of Business Activity in the Puget Sound Area.* Bervard Nichols 483

CRUM, WILLIAM L., FENNELLY, JOHN F., AND SELTZER, LAWRENCE H. *Fiscal Planning for Total War.* Alfred G. Buehler 479

FABRICANT, SOLOMON. *Employment in Manufacturing, 1899–1939.* Daniel Creamer 481

FENNELLY, JOHN F., see Crum, William L.

HEANEY, NORMAN S. *Public Trusteeship.* C. Herman Pritchett. . . . 487

JACOBY, NEIL H., see Saulnier, Raymond J.

LANDSBERG, HANS H., see Barger, Harold

MAGILL, ROSWELL. *The Impact of Federal Taxes.* Richard A. Musgrave . 476

Continued on next page

REVIEWS—*Continued*

National Bureau of Economic Research, Conference on Research in Income and Wealth. *Income Size Distributions in the United States, Part I.* Daniel Creamer 475

Rostas, L., see Shirras, G. Findlay

Saulnier, Raymond J., and Jacoby, Neil H. *Accounts Receivable Financing.* Robert D. Ayars 480

Seltzer, Lawrence H., see Crum, William L.

Shirras, G. Findlay, and Rostas, L. *The Burden of British Taxation.* Tipton R. Snavely 478

Sturmthal, Adolf. *The Tragedy of European Labor, 1918–1939.* Robert I. Crane 488

Tax Institute. *Tax Yields: 1941.* Henry S. Bloch 477

Tomasevich, Jozo. *International Agreements on Conservation of Marine Resources.* Reginald H. Fiedler 486

Walker, Helen M. *Elementary Statistical Methods.* Laverne K. Burke . 489

Wilks, S. S. *Mathematical Statistics.* Abraham Wald 491

Zapoleon, Marguerite W. *Community Occupational Surveys.* Harold F. Clark 490

Publications Received 493

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 224 DECEMBER, 1943 *Volume 38*

AN APPRAISAL OF THE U. S. BUREAU OF LABOR STATISTICS COST OF LIVING INDEX*

BY A SPECIAL COMMITTEE† OF THE AMERICAN STATISTICAL ASSOCIATION

FOREWORD

THE COST OF LIVING INDEX published by the Bureau of Labor Statistics has been one of the important statistical tools of the country for a quarter of a century. Designed to measure changes in the prices paid for goods and services by city consumers of moderate income, it is widely used by statisticians, economists and government officials. It has often been an important factor in wage negotiations, and has served to throw light upon the effects of fiscal and other governmental policies. In technical statistical work it serves as a "deflator" to be used in computing "real" income and expenditures.

---

* A supplementary report, referred to as the Appendix in this general summary, will appear in print later. This supplementary report will present the statistical evidence which the committee has assembled, together with the committee's technical recommendations.

† The committee consisted of the following members:

Frederick C. Mills, *chairman*, Professor of Economics and Statistics at Columbia University and a Director of the National Bureau of Economic Research; former President of the American Statistical Association and of the American Economic Association; author of *The Behavior of Prices*, *Economic Tendencies in the United States*, and *Prices in Recession and Recovery;*

E. Wight Bakke, Professor of Economics at Yale University; Hearing Officer, representing the public on national cases for the War Labor Board; author of *The Unemployed Man*, and *Citizens Without Work;*

Reavis Cox, Professor of Marketing, at the University of Pennsylvania (on leave); Director of Research Projects for the Retail Credit Institute of America; Editor-in-Chief of the *Journal of Marketing;* author of *Competition in the American Tobacco Industry*, and *Marketing of Textiles;*

Margaret G. Reid, Professor of Economics, at Iowa State College; author of *Economics of Household Production*, *Consumers and the Market*, and *Food for Poeple;*

Theodore W. Schultz, Professor of Economics at the University of Chicago, and Consultant to the Board of Governors of the Federal Reserve System; author of *Vanishing Farm Markets and Our World Trade*, and *Farm Prices for Food Production;*

Samuel Stratton, President of Middlebury College; co-author of *The Economics of the Iron and Steel Industry; Problems in Corporation Finance*, and *Financial Instruments and Institutions.*

Appointed by the committee to serve as its technical staff were the following:

Dorothy S. Brady, Senior Statistician, Bureau of Human Nutrition and Home Economics, U. S. Department of Agriculture; author of miscellaneous articles and reports particularly in the field of family income and expenditure;

Solomon Fabricant, member of the research staff of the National Bureau of Economic Research; Head Economic Analyst, War Production Board; author of *Capital Consumption and Adjustment* and *Employment in Manufacture.*

As wartime strains and distortions have spread through the American economy, criticisms of the index have developed. In particular, its use as a guide in wage adjustments has been widely questioned. Many critics express doubt of its accuracy as a measure of the extent to which the country has experienced inflation under the stimulus of huge governmental expenditures. Much of what these critics say takes the form of declarations that the index does not provide an accurate measure of what is happening to consumers' expenditures.

This loss of confidence in a basic and generally accepted statistical measure has occasioned much concern in professional and governmental circles. Accordingly, Dr. E. A. Goldenweiser, President of the American Statistical Association, at the request of the Bureau of Labor Statistics, transmitted through the Secretary of Labor, appointed this committee to review and appraise the cost of living index with reference both to its construction and to its uses. Having named the committee, he left it completely free and independent to pursue its investigations in such directions as seemed appropriate to it.

In our review we have canvassed interested labor organizations, employers' associations, and consumer groups. Evidence on changes in living costs has been submitted to us by representatives of these groups and by governmental agencies. Special field studies have been initiated by the committee and conducted by agents of the Bureau of Labor Statistics under the supervision of committee representatives.

The evidence we have assembled, detailed explanations of our findings and certain technical recommendations will be given in a supplementary report. We here present our conclusions in a condensed general summary followed by our answers to questions that pose the major issues.

GENERAL SUMMARY OF CONCLUSIONS

The principal conclusions to which we have come are: First, that within the limitations established for it, the cost of living index provides a trustworthy measure of changes in the prices paid by consumers for goods and services. Second, that many of the difficulties and doubts which have arisen concerning the index have their origins in attempts to use it uncritically for purposes to which it is not adapted.

The phrase "cost of living" is ambiguous. Some of the difficulties the use of this index raises grow out of this fact. In every-day speech families are apt to think of their cost of living as the total amount they spend for consumer goods and services. Several different factors may cause such expenditures to change. Thus, they may change because the unit prices of goods and services rise or fall. Again, they may change because families are forced by circumstances beyond their control to alter their manner of living, as when the exigencies of war make some goods unavailable. Yet again, they may change because families have experienced an increase in income and can afford to buy more goods or better goods.

Changes caused by alterations in manner of living and in income have been particularly important during the war. This is true because of the very great migrations of labor and shifts in occupation characteristic of the period. Because of shifts from lower to higher-paid jobs, rises in wage rates and extension of the work week, many individual and family incomes have advanced more rapidly than living costs. Studies by the Department of Commerce indicate that the annual average compensation of wage earners and salaried workers in private non-agricultural industries increased 37 per cent between 1939 and 1942 (from $1268 to $1733). Living costs, as measured by the Bureau of Labor Statistics index, increased 17 per cent in the same period.

As used in technical statistical parlance, the term "cost of living" has applied only to the first of the factors which determine family expenditures, that is, to unit prices. The cost of living index has undertaken to measure only this one factor. Any use made of the index must keep this limitation in mind. Many of the difficulties which critics of the index have encountered arise from forgetting that it is designed to measure only changes in retail prices and trying to use it for purposes to which it is not suited unless it is supplemented by a number of additional statistical measures.

After a careful investigation of the relevant evidence we believe that the index records with a satisfactory degree of accuracy what has happened to unit prices paid for consumers' goods by city families of moderate income.

We have been much impressed during our investigation by the integrity, alertness, intelligence and technical competence of the Bureau's staff. We feel that, on the whole, working under very trying circumstances, they are doing an excellent job of carrying out the task assigned to them.

MAJOR ISSUES BEFORE THE COMMITTEE

1. *Are the prices reported to the Bureau those actually paid by customers of reporting retail stores?*

The Bureau obtains its prices through field agents who interview the owners or managers of the stores they visit. They do not actually purchase the merchandise they price. Prices for chain stores are obtained, as a rule, not in the stores themselves, but from the central offices. This method of collecting prices has led to the criticism that the quotations reported will be higher or lower than actual prices according to whether it is to the interest of the retailer to appear to be holding prices above or below the true market. During a depression period, when there is strong business pressure on retailers to avoid price cuts, a tendency for retailers to report prices higher than those they actually charge might be expected. On the other hand, during the present war, when the regulatory pressure is to hold prices below the levels they would reach in a free market, there may be a tendency for retailers to report prices lower than those they actually charge.

Answer: We have tested reported food price quotations through actual purchases at 159 independent stores and 37 chain stores in 17 cities.[1] Stores visited were selected from those reporting to the Bureau. Our buyers made these purchases as actual customers, not as Bureau or committee agents. Eleven hundred and twenty-seven purchase prices were compared with the same number of prices reported by the stores for the same dates.

862 purchase prices (76 per cent) were identical with reported prices;
90 purchase prices (8 per cent) were lower than reported prices;
175 purchase prices (16 per cent) were higher than reported prices.

We find a slight tendency for prices reported to the Bureau of Labor Statistics to understate prices actually charged customers. In its net effect this understatement involved a difference of not more than 1.8 per cent of the total cost of the foods priced.

To check the accuracy of food prices reported to the Bureau by central offices of food chains, agents of the Bureau, under the supervision of this committee, visited 200 individual chain stores, distributed among 6 cities.[2] Price reports to the number of 10,366 were obtained, for comparison with those reported by central offices for the same date.

8,023 prices in individual stores (78 per cent) were identical with those reported by central offices;
970 prices in individual stores (9 per cent) were lower than those reported by central offices;

---

[1] A test on May 18, 1943, covered Atlanta, Boston, Chicago, Cleveland, Dallas, Denver, New York, and San Francisco. A test on June 15, 1943, covered Atlanta, Detroit, Houston, Milwaukee, Newark, New York, Philadelphia, Salt Lake City, San Francisco, Seattle, Scranton, and Washington.

[2] The test, made on June 15, 1943, covered Atlanta, Detroit, Milwaukee, San Francisco, Scranton, and Washington.

1,373 prices in individual stores (13 per cent) were higher than those reported by central offices.

On balance, there is a slight tendency for prices reported by central offices to understate the prices charged in chain store outlets. In its net effect this understatement involved a difference of 0.6 per cent of the total cost of the foods priced.

We believe that these checks confirm the essential accuracy of the food prices reported to the Bureau. The observed discrepancies indicate a margin of reporting error of less than 2 per cent.

Our field tests were confined to food prices because these have been most frequently questioned in our hearings. However, members of the committee have accompanied the Bureau's agents and have observed the collection of the prices of other consumer goods. We believe these to be substantially accurate.

It should be noted that the obtaining of reported prices which at any one time are lower than the true prices does not necessarily reduce the accuracy of the index as a measure of changes over a period of time. If merchants since the beginning of the war have consistently under-reported their prices, the trend of the index would not be affected. In any event, since the understatement in June 1943 was slight, it may be inferred that even if merchants tended to understate their prices somewhat more than before, the effects upon the index would be negligible.

2. *Are the retail stores included in the Bureau's sample representative of those patronized by wage earners and lower-salaried workers in metropolitan areas covered by the Bureau's compilations?*

The Bureau selects the stores from which it obtains prices on the basis of criteria which indicate where wage earners and lower-salaried workers in the city buy goods and services. The specific stores from which its agents obtain price quotations are necessarily changed over a long period but they are changed slowly. Some users of the index believe that migrations of labor during the war have changed the purchasing habits of the workers so drastically that the group of stores priced by the Bureau is no longer representative. They stress especially the belief that the changes in shopping habits have had the effect of transferring patronage from low-priced to high-priced stores. One example often given of this is the possible transfer of trade from city stores to suburban stores. Another illustration is the possible transfer of trade from chain to independent stores.

Answer: The only adequate way of answering this question would have been to ask a representative sample of workers where they made their purchases. The Bureau, in selecting its sample of stores, considers many factors in order to assure itself that the stores selected are in neighborhoods where the income groups concerned purchase goods. We are satisfied that the Bureau's criteria are the best that can be devised without a new family survey. The weight given to chain and independent stores is based on comparative volume of business with respect to all income groups and not merely those included in the index. It has recently been revised to take account of wartime shifts in buying. If low income purchasers have shifted their marketing from one type of store to another, the weight given to each in the index may not truly reflect their comparative volume of business with respect to this group.

Since the elaborate survey necessary to answer this question by actual study of the purchasing habits of workers and other low income groups could not be made by this committee, we have undertaken two checks which throw indirect light on this problem.[3] Food prices were obtained in June 1943 from a number of independent retail stores in addition to those normally surveyed by the Bureau. One hundred and ninety-five of these were located within the limits of six cities[4] and 145 were located in the suburban areas of four cities in which there had been marked wartime expansion.[5] The latter sample was selected to ascertain whether prices in outlying areas in which there had been considerable population increases were higher than prices in the central cities. The criteria of selection employed were those used by the Bureau in obtaining its present sample of reporting stores. For each of 78 food products priced the average price in independent stores in the regular sample in a given city was compared with the average price in the added sample of independent stores in outlying districts of the city.

Four hundred and fifty-three[6] comparisons were made in testing the regular samples against samples drawn from within the limits of the same cities.

37 average prices from the check sample (8 per cent) were identical with those from the regular sample;
269 average prices from the check sample (59 per cent) were lower than those from the regular sample;
147 average prices from the check sample (33 per cent) were higher than those from the regular sample.

---

[3] The methods used in making these checks and their limitations are described in the Appendix.

[4] Atlanta, Detroit, Milwaukee, San Francisco, Scranton, and Washington.

[5] Detroit, Milwaukee, San Francisco, and Washington.

[6] Not all of the 78 foods were priced in each city.

The net effect of these discrepancies was to make the total cost of the food priced in the added sample of stores within city limits 1.7 per cent less than the cost of the same foods in the regularly reporting stores in the same cities. Statistical tests described in the Appendix show that this difference is not large enough to be significant.

Three hundred and five comparisons[6] were made in testing the regular sample against samples drawn from outlying areas of the same cities.

30 average prices in the outlying areas (10 per cent) were identical with those from the regular sample;
165 average prices in the outlying areas (54 per cent) were lower than those from the regular sample;
110 average prices in the outlying areas (36 per cent) were higher than those from the regular sample.

The net effect of these variations was to make the total cost of foods priced in the sample of stores in outlying areas 1.7 per cent less than the cost of the same foods in the regularly reporting stores in the cities proper.

We conclude that food price reports from the retail stores included in the regular Bureau sample are representative of prices prevailing in stores patronized by wage earners and lower-salaried workers in the cities covered by the tests made and in suburban sections of the metropolitan areas tested. In particular, we find in this evidence no support for the opinion that prices in independent stores in suburban districts are consistently higher than prices in independent stores in neighboring large cities.

These comparisons rest on samples of independent stores in representative metropolitan districts. We believe that the conclusions drawn are generally valid for such areas in the country at large. Exceptions may be found in specific boom areas. Smaller cities and towns are not included in the Bureau's cost of living index.

The tests here reported relate to price levels on a given date. On the committee's recommendation similar tests are being made for subsequent months in order that the regular sample and the added samples may be compared with reference to price changes over a period of time.

3. *To what extent is the Bureau's index of living costs affected by the omission of black market prices?*

Wartime ration and price controls have led to a widespread belief that much retail trade is now being done at prices far above officially established ceilings through both regular and irregular channels of distribution. If this were true, and if above-ceiling prices were not adequately represented in the Bureau's compilations, the cost of living index would be biased downward.

Answer: We distinguish between black markets proper, in which rationed goods are sold without exchange of ration coupons, and retail stores in which some goods are sold at above-ceiling prices, with surrender of ration coupons by customers. We have made no systematic investigation of markets of the first type. The questioning of informed witnesses indicates that the families to which this index relates do not purchase important quantities of goods in such markets. The Bureau does not attempt to obtain price quotations in these markets. It is the opinion of the committee that it should not do so. *Transactions in rationed goods effected without exchange of ration coupons should not be represented in an official index of living costs.*

The Bureau does obtain above-ceiling prices charged in ordinary retail outlets, with customary exchange of ration coupons, and these prices are included in the index. Among the prices reported for April and May 1943 for pork products, 1,981 out of a total of 8,707 (23 per cent) were above ceilings. It is not likely, however, that all such cases are reported to the Bureau. The net difference between reported prices and the purchase prices recorded in our field tests is an indication of the possible magnitude of the effect of the failure to report all above-ceiling prices. We conclude that there is probably an error in the index because of the under-reporting of above-ceiling prices by ordinary retail outlets, but that for food products this error does not exceed 2 per cent.

4. *Does the index of living costs measure hidden price increases due to quality deterioration?*

Although the formal quotation for an item may remain unchanged, its true price may be raised by reducing the quality offered. Thus, a shirt quoted at $1.95 throughout the war may have increased substantially in price if the quality has deteriorated. Critics of the index argue that this sort of hidden increase in prices has been very large and is not reflected in the index.

Answer: The Bureau of Labor Statistics seeks to hold quality factors constant by pricing goods with specified characteristics. For many types of consumer goods it is able to do so today. For other goods and services quality deterioration is an inevitable accompaniment of the war effort. The lessened comfort of traveling, reduced delivery and laundry services, the use of cuts of meat of lower grade, are examples of quality deterioration that we accept as necessary under war conditions.

With limited exceptions these quality reductions are unavoidable. This does not mean that all consumers are today buying goods of lower quality. Many workers whose incomes have risen during the war are buying clothing and food of higher quality than they could afford in the past. These come at higher prices, and mean to the buyer a higher level of family expenditures. To the degree that higher quality goes with the higher prices this is a change in the level of consumption, not an increase in the cost of living.

With reference to the cost of living index, our concern is with the price equivalent of quality change. *We believe that consumers' goods and services, in the aggregate, have since 1939 suffered some loss of quality that is not reflected in reported prices. No dollar value can be put on this loss. In large part it is an intangible and unmeasurable element of the war. Consumers cannot be compensated for it. The cost of living index takes incomplete account of it.*

5. *Are the rents reported to the Bureau those actually paid by tenants?*

Some users of the index have been under the impression that rentals are still obtained from rental agents, who are under pressure from the rent control authorities, and that the rent reports are, therefore, biased downward.

Answer: We have made no field check on this point, but we have examined the Bureau's procedure. We find that all rental reports are now obtained directly from tenants by agents who visit their homes. We believe that the rents thus reported are free from bias and may be accepted as accurate.

6. *Are the housing facilities included in the sample representative of those occupied by wage earners and lower-salaried workers in large cities?*

Answer: The Bureau's housing sample prior to September 1942 was made up of houses and apartments handled for the owners by rental management firms. This sample was losing its representative character under conditions of housing scarcity and rent control. Between September 1942 and June 1943 a complete new sample was drawn in each city covered. The total number of family dwelling units included is 51,500.

In our opinion the new samples are satisfactory representations of family dwellings in the cities represented in the general index. We do not believe that these 34 cities adequately represent the diversity of urban housing conditions in the country as a whole. We recommend that advantage be taken of rental surveys now in process or being planned by the Bureau and the Office of Price Administration to increase the coverage of the index.

We note one other limitation on the coverage of the Bureau's sample. This includes family dwelling units only. Under wartime conditions single rooms and portions of former family dwelling units have become more important as living quarters for workers. Evidence presented to the committee indicates that rental rates for individual rooms and for parts of family dwelling units have risen more rapidly in many congested war centers than have rentals for whole family dwelling units. We note, however, that the amount spent on room rentals in the spring of 1942 by urban families of two or more persons amounted to only about 1.9 per cent of all expenditures on housing by such families (*Study of Spending and Saving in Wartime*, a pending publication of the Bureau of Labor Statistics). Today such expenditures may be in the neighborhood of 2 per cent of the total amount spent on housing by urban families of moderate income. The inclusion of these expenditures in the Bureau's rent index, even at levels well above those of 1942, would affect the total cost of living index only slightly.

7. *Does the Bureau's rent index accurately measure recent changes in the cost of housing to wage earners and lower-salaried workers living in large cities?*

Answer: *We believe that for the areas covered by the Bureau's sample the index of rents provides an acceptable measure of changes in the cost of housing for urban workers who live in the same communities and in residences of the same general types as those occupied before the defense and war effort.* We point out, however:

(a) There has been deterioration in the quality of housing in certain regions. For example, houses on the average are older, and a larger proportion of those included in the index are in need of repairs and renovations if pre-war standards are to be maintained. This deterioration is in some degree unavoidable and in some part unmeasurable. It constitutes a loss that is not taken account of in the index of living costs. However, the committee would point out that certain directly measurable changes, such as the discontinuance of "free rent" for certain periods of the year or the shift of payments for water or heat from landlord to tenant are recorded by the Bureau as rent increases.

(b) Workers who have moved from low-rent to high-rent areas and workers who have moved from family dwelling units to rooms in lodging houses or homes have experienced substantial advances in housing costs. These advances are not measured by the rent index of the Bureau of Labor Statistics. These changes are illustrations of the changes in manner of living referred to above.

Although they may have an important effect upon family expenditures they cannot be measured by an index designed to record changes in unit prices over a period of time.

8. *Are the goods and services priced now typical of those purchased by wage earners and lower-salaried workers in large cities?*

An index of the cost of living, if it is to be meaningful, must measure changes occurring in the prices of what is actually consumed. The committee has given attention, therefore, to the methods by which the Bureau has adjusted the elements of the index to changes in consumption.

Answer: The basis for the selection of commodities and services included in the index is a study of consumer expenditures relating to the years 1934–36. The standards represented in that selection related to the consumption of urban families supported by wages or clerical salaries with incomes averaging about $1,500. These standards were somewhat higher than the level of expenditures of this group in the thirties because the sample did not include families receiving relief of any kind or families with incomes under $500. With the rising incomes of 1939 and succeeding years, wage earners' expenditures probably came closer to the standards on which the cost of living index is based. As of 1940 the commodities and services included in the index were probably more representative of actual consumption by urban workers than they had been five years before. Since 1941 rationing and the disappearance of commodities have altered the kinds and quantities of goods and services consumers can buy. Corresponding changes have been made in the composition of the index of living costs.[7] Durable metal goods have been virtually withdrawn. The current food index was adjusted in March 1943 on the basis of the Department of Agriculture's estimates of expected consumption of different kinds of food in 1943. Other elements of the index have been altered to accord with rationing orders. There is no completely satisfactory solution of the problem faced by makers of a cost of living index when the character of consumption is changing. We believe that the adjustments made by the Bureau represent a reasonable compromise between the desire to reflect current consumption and the desire to maintain an index that measures with accuracy average changes in the unit prices of consumer goods and services.

We would add that changes in the kinds and quantities of goods consumed by average families in the United States have not as yet been very great. Mass consumption is slow to change. Purchases of new durable goods are, of course, smaller, but such goods constitute in normal times only some 10 per cent of consumer outlay for families with moderate incomes. For wage earners and salaried workers in the country at large, the distribution of expenditures for food, clothing, and rent has not been substantially modified.[8] The index of living costs, it must be remembered, relates to changes in the average situation of city wage earners and clerical workers, not to the experience of particular families.

---

[7] See Appendix. See also "Bureau of Labor Statistics Cost-of-Living Index in Wartime," by Faith M. Williams, in the July 1943 *Monthly Labor Review.*

9. *Does the index of living costs measure changes in average family expenditures on consumer goods and services?*

This question reflects the popular interpretation of the phrase "cost of living" as representing the total amount families spend for goods and services.

Answer: No. The index is designed to measure only those changes in family expenditures resulting from changes in unit prices. It is not designed to measure alterations in family expenditures resulting from changes in manner of living, whether the changes are due to choice or necessity.[9] Such alterations can be measured only by other statistical devices, such as direct inquiry into actual family incomes and expenditures or the index of consumer expenditures now being published by the Department of Commerce. The distinction between changes in family expenditures due to price movements and those due to shifts in the way of living is of special concern today when the pattern of living is being altered for a great many consumers. Rising incomes are giving many families opportunities to live on higher levels than they have known before.

10. *Does the Bureau's index provide an acceptable approximation to recent changes in the cost or living for urban workers?*

Answer: Yes.

The index shows the following changes in the cost of living in the United States:

from March 1943 to August 1943	+ 0.3 per cent
from September 1942 to August 1943	+ 4.6 per cent
from January 1941 to August 1943	+22.2 per cent
from August 1939 to August 1943	+24.9 per cent

---

[8] See the comparative distributions of consumer expenditures from 1939 to 1942, *Survey of Current Business*, March 1943.

[9] Certain forced changes are taken into account by recently introduced methods of treating the difference in price between a higher-priced article and one no longer available as a price increase.

The various tests we have made indicate that these measurements define with satisfactory accuracy changes in the prices of the goods and services purchased by urban workers.

With this statement we repeat two necessary explanations and two necessary qualifications.

(a) The index relates to living costs for wage earners and lower-salaried workers in large cities.

(b) The index is primarily a measure of changes in the average unit prices of consumer goods and services. It does not measure changes in family expenditures resulting from changes in income or in the place or manner of living.[10] The general effect of the shifts in incomes and in ways of living that have occurred since 1939 has been to advance total family expenditures at a rate greater than that shown by the Bureau's index of living costs.

(c) The index does not measure the full effects of quality deterioration in certain consumer goods and services and in housing.

(d) The kinds of goods and services used by the average family have changed somewhat under the impact of war and will continue to change. The Bureau has made changes in the list of goods that it prices and has altered the relative importance of component items to take account of the disappearance of goods and changes in consumption resulting from rationing. These shifts are reflected in the index. Their effect on the actual course of the index has not been great. Checks made by us show that price changes have dominated the movements of the index in recent years.

We have said that the index provides an acceptable approximation to recent changes in living costs. *We believe that as a measure of price changes affecting urban workers in large cities it is a good approximation.*

11. *Is the index an adequate instrument for use in matters of public policy?*

Those charged with the formulation of public policy and its implementation must frequently make use of an index of living costs. The President, Congressional committees, and public agencies such as the War Labor Board, the Social Security Board, the Office of Price Administration, the Treasury, Bureau of the Budget, State Department, and others must consider the economic experience of citizens as reflected in the index in formulating and in evaluating the effect of their several policies.

---

[10] Many such changes have come with recent migrations of labor and occupational shifts. Because occupational shifts have, in general been from lower to higher pay industries and from lower to higher paid jobs within industries, personal and family incomes have advanced more rapidly than average wage rates and much more rapidly than living costs. The Department of Commerce estimates that the average annual compensation of employees (wage earners and salaried workers combined) in private nonagricultural industries increased 37 per cent between 1939 and 1942 (from $1,268 to $1,733). Living costs increased 17 per cent in the same period.

Answer: We have indicated our conviction that the index is a satisfactory instrument for measuring what it attempts to measure—average movements in the retail prices of goods and services purchased by city workers. If public policy is in need of a guide on this point, the index is a satisfactory one.

In general it can be said, therefore, that the greater the dependence of public policy upon a measurement of what is happening nationally to prices of consumers' goods, the more useful the index is. On the other hand, the greater the dependence of public policy upon a measurement of what is happening to prices in particular areas, for particular population groups, and for particular commodities not covered by the Bureau, the less useful it becomes.

The value of the index for certain uses is greatly enhanced if it is used as one among several measures rather than alone. For some problems it should be used as a first approximation to the truth, to be corrected by knowledge of the particular situation. A thermometer reading is useful to the physician if used in conjunction with other indicators of the health of the patient and as a guide to further diagnosis.

We cite three examples of public use of the index to illustrate our point:

(a) *Use of the index to measure the degree of inflation.*

If we define inflation as a situation in which the aggregate value of all goods and services exchanged in an economy increases more rapidly than the physical volume of these goods and services, a cost of living index is an obviously relevant measuring tool. The Bureau of Labor Statistics index, however, does not cover all consuming groups in our economy. It relates only to the retail prices of consumers' goods purchased by families of moderate income in large cities. Prices in primary markets and in various wholesale markets are also important indicators of inflation. For this reason we do not believe that the Bureau's cost of living index should be used by itself as a measure of the extent of inflation. It has great value as one among a group of measurements available to administrators of an anti-inflation program.

(b) *Use of the index in the control of retail prices.*

The index is a useful instrument in retail price control. Its use would be greater if coverage were broadened to include prices of more commodities. However, it should be remembered that the commodities selected for measurement are chosen not only because they are important in the customary purchases of consumers, but also because their price movements are like those of related commodities. If the present national and city averages were supplemented by regional measures and by more detailed index numbers for commodity groups, their usefulness would be greatly increased. Effective price control cannot be based solely upon averages comprehending great diversity.

(c) *Use of the index as an instrument in making wage and salary adjustments.*

On the general question of policy involved in deciding whether wage rates and salaries should in times of peace be adjusted in accordance with movements of an index of living costs the committee does not speak. However, we have been asked to appraise the Bureau's index with reference to the uses to which it is now put. If the index is to be appropriately used in this connection, several facts must be kept in mind.

The general index is a national average, and relates to large cities only. There is considerable variation in the level of living costs in different parts of the country, and in the degree of change in living costs in different regions. There are industrial variations in living costs. Costs in coal towns and textile towns are affected by special factors and have distinctive movements. Living cost changes vary also with income levels. Recent advances have been somewhat greater for workers at lower income levels because foods, which have risen sharply in price, are of greater importance in the budgets of low-income families.[11] Such variations are characteristic of economic changes in a system as diverse as that of the United States.

These facts have relevancy for several possible uses to which the index may be put in connection with wage and salary adjustments. In the first place, *in the application of general wage policies of national scope it is proper and necessary to utilize a measure that averages these varying movements, as a national index does. If wages and salaries of workers in particular regions, in particular industries and at particular income levels are to be adjusted to living cost changes affecting them specifically, indexes of living costs adapted to special circumstances will be required.* The construction of adequate regional and other special purpose index numbers of living costs would involve broader sampling and heavier expenditures by the Bureau of Labor Statistics.

---

[11] Furthermore, the actual family expenditures of low income groups have probably risen more than those of higher income groups because consumers at low income levels have not been in position to economize by purchasing goods of lower grades than those to which they have been accustomed in the past.

The use of the index in such specific wage adjustments is not, of course, the one which is currently most significant. Wage rates cannot be contractually geared to the movements in the cost of living index under the Federal Government's present stabilization policy if to do so would result in wage increases inconsistent with that policy. In the Government's stabilization program the cost of living index has two uses: first, to explain the 15 per cent increase in wage rates permitted by the "Little Steel" formula, and second, to give an indication of the pressure placed upon present wage rates by workers' experience in meeting the rising cost of living. We have found no evidence that would invalidate the index for the first purpose. For the second purpose, however, that of estimating the pressure on present wage rates and the present wage adjustment formula by workers' experience with living costs, a national average of living costs is not satisfactory. The reason for this does not arise from any inaccuracy, but from the fact that it is an average. The experience from which dissatisfaction comes is not that of the statistical average man. It is the experience of the individual worker or group of workers involved.[12]

The fact that the index is one of prices and not of consumers' actual expenditures should provide further warning. The difference between the movements of prices and changes in actual expenditures is significant in times of great mobility, changes in the pattern of living, commodity and housing shortages, quality deterioration, and increased taxes and bond purchases. Since a price index cannot reflect completely these facts of consumers' experience, it cannot fully measure what workers mean by changes in their cost of living. The dissatisfaction or satisfaction of workers arises from their own and not the statistician's conception of the "cost of of living" in relation to their own and not the average income.

One further point needs emphasis. A decline in the plane of living cannot be prevented in an economy that absorbs a large part of the country's products for military purposes. When the total available supply of consumers' goods declines, the real income and the plane of living of the average citizen must decline. If the wages and salaries of any group or the prices received by any group are geared to a cost of living index, making it possible for them to purchase the same volume of goods as before, the consequences are clear. The remainder of the population whose income is not geared to the index must suffer a greater decline in their plane of living than would be expected in view of the reduction in the available supply of consumers' goods. It is not intelligent conservation of the nation's human resources, of course, to permit a reduction in the plane of living of those living at or near subsistence levels. But for the population at large war means reduced consumption.

---

[12] It should be pointed out also that the worker's dissatisfaction with living costs is modified somewhat by his earnings and that changes in total earnings from whatever causes are important in this connection as well as changes arising solely from the alteration of wage rates.

The use of a cost of living formula to adjust the income of one group of citizens will not prevent reduction of consumption for citizens generally. It can result only in shifting a larger share of the burden of that reduction onto the shoulders of those whose income is not geared to the formula. The preferable alternative is one that combines rigorous and extensive taxation, government borrowing from individuals, rationing and price control as means of restricting consumption, keeping inflationary forces in check and insuring an equitable sharing of the burdens of war.

RECOMMENDATIONS TO THE BUREAU OF LABOR STATISTICS

In the course of its hearings and work the committee has reviewed the many uses the Bureau of Labor Statistics must take into account in making its cost of living indexes more effective during the war and in the coming period of reconstruction. The following recommendations are based upon this review. They are made with the understanding that proper consideration will be given to the costs of the studies proposed, as well as to the needs of administrators and legislators for the information these studies will yield. The Bureau and the appropriate Congressional committees will weigh the expected benefits against these costs.

With this proviso, the committee recommends:

1. That the Bureau of Labor Statistics provide *separate regional* indexes of changes in city living costs. The economic differences between the several large geographic areas of the United States make it essential that separate indexes of the cost of living for various regions be prepared.
2. That the Bureau expand its indexes of the cost of living in small cities so that they will be as representative of these cities as the large city indexes are for large cities.
3. That the Bureau cooperate with other interested governmental agencies to collect the information necessary and to prepare an

index of the cost of living for *non-farm* communities in the United States.

4. That the Bureau take steps immediately to prepare and publish cost of living indexes for:
   a. Communities of different sizes.
   b. Certain specialized industrial areas, such as those in which coal mining, textiles, etc., are concentrated.
5. That the Bureau immediately expand its work on the development of measures of geographic differences in cost of living; and that analyses of these differences be continuously carried on.

To carry out the foregoing recommendations, it will be necessary for the Bureau immediately to expand its samples of large and small cities, and employ existing information and collect additional data to develop weights which will reflect the geographic, industrial, and economic factors involved in the proposed indexes.

6. That frequent *small-sample studies of family expenditures and incomes* be conducted, and that once in 5 years a comprehensive study be made to provide data on local differences in income and consumption habits among occupational groups in the United States. Furthermore, that adequate analyses and integration of these studies and existing data be planned and regularly provided for. Since these kinds of studies provide the weights which reflect the changes in consumption habits and are important in keeping the cost of living indexes in line with current economic development, it is essential that they be made periodically.
7. That the Bureau undertake analyses of the shifting importance of the various items included in family budgets and the degree to which these items are correlated with other items in price movement. To keep sufficiently abreast of these kinds of changes, the committee feels that a comprehensive study of the inter-relationships of price changes and of the correlation between price changes and other factors (size of community, quality and season) should be made.
8. That the Bureau *publish average prices* for the principal commodities and services in each of the major groups of consumption and that as rapidly as the validity of the samples and averages of sub-groups of these goods and services can be assured, that indexes and average prices for these groups also be published. For general index purposes, some of the Bureau of Labor Statistics samples (of cities, goods and services) might be reduced; the pressure for increasing detail on the part of those

using these indexes, however, requires larger samples for most phases of the work. At the present time—except for foods, fuels, and certain other commodities—the Bureau of Labor Statistics is not warranted in publishing basic data on average prices by cities because the samples have in general been considered as inadequate to provide more than stable relatives for large aggregates of commodities and services.

In periods of rapid changes in consumption habits, the specifications of goods and services purchased by consumers, as well as the lists of stores from which consumers buy, must be reviewed at frequent intervals if the cost of living index is to be maintained as a dependable instrument.

9. That the Bureau pay continuous attention through special studies to changes in the *specific consumer goods and services* which it prices, continuing the survey of wartime prices initiated in the spring of 1943.
10. That the Bureau conduct frequent field surveys to determine the necessity of changing the number of outlets from which prices are obtained, or of shifting from outlets in one area to those in another, in order to maintain adequate samples for the indexes, as well as for the average prices.
11. That the Bureau prepare for publication an historical summary of the development of the concept of the cost of living index and of the purposes for which that index is used. The Bureau of Labor Statistics index of the cost of living has been developed with gradual improvements in methodology since the introduction of the index of food costs in 1903 by the Commissioner of Labor Statistics. The report herein recommended should, in addition, describe the expansion of uses to which the index has been put, and provide an explanation of the ideal form of index which would serve these several purposes.

FREDERICK C. MILLS, *Chairman*
E. WIGHT BAKKE
REAVIS COX
MARGARET G. REID
THEODORE W. SCHULTZ
SAMUEL S. STRATTON

HOUSING SCALES FOR RURAL PENNSYLVANIA

By Howard R. Cottam
The Pennsylvania State College

Evaluations of housing, whether of a slum area or of an individual dwelling, are always based upon certain criteria. Too often the criteria are ill chosen and inconsistently weighted. In this article it is assumed that the comparative usefulness of various criteria can be empirically determined and weighted accordingly. What follows is a description of the reliability and the validity of three scales of varying length for measuring rural housing in Pennsylvania. The scales were constructed from data collected in 1941 through interviews with 517 rural householders[1] concerning 57 housing items.

Housing, for purposes of this study, was defined broadly to include the house structure together with the equipment and the site. More precisely, housing was defined in terms of the 57 specific component items, although at this stage the relative importance of each had not been determined. One object of preparing a scale was, of course, to permit measurement in terms by which housing could be defined more explicitly, that is, operationally.

In constructing the scales two preliminary subjective evaluations were made by the author: first, the selection of the 57 specific items; and second, the designation of categories for each of the 57 items and their arrangement on a continuum from poorest to best housing (i.e., poorest or best as judged subjectively in terms of desirability in Pennsylvania culture). Four distinct methods of assigning weights to the categories of each item were then tried. After testing the reliability and validity of the weights finally adopted, the possibility of reducing the number of items without sacrificing too much reliability and validity was explored and three abbreviated scales comprised of 10, 20, and 30 items were constructed.

The first method involved the assigning of *a priori* weights to each category of each of the 57 component items by the author. A second method used as weights, values which took account of the frequency of occurrence of each category of each component item; that is, the percentages of families possessing each category of each component were converted into standard-deviation units by reference to an area table. A third method was based upon calculated values which indicated the degree of probability that each item was diagnostic of total housing, that is, critical ratios of differences between high-scoring and low-scoring classes of families. A fourth method determined weights for each category of each component item through a series of successive approximations in which *a priori* weights were successively revised by assigning as an item weight the average score of all other families possessing that item. Each of these methods is described more fully below. Let *A*, *B*, *C* and *D* designate the four methods in the order described above.

---

[1] For a statement of the sampling procedure see *Housing and Attitudes Toward Housing in Rural Pennsylvania*, Pennsylvania Agricultural Experiment Station Bulletin 436, and *Measurement of Housing and Attitudes Toward Housing in Rural Pennsylvania*, Pennsylvania Agricultural Experiment Station Technical Paper No. 1149 (mimeographed).

Method A of assigning weights to each category of each housing item was admittedly subjective. It rested upon preconceptions of the author as to the relative importance in rural Pennsylvania culture of each characteristic (each type or category) of each of the 57 component items and upon the assumption that each component item should be given approximately equal weight in the total score. More explicitly, the author arbitrarily decided that the presence of an electric sweeper (one category of one of the 57 component items) warranted a higher score than the presence of a hand sweeper (mechanical but not power-driven), and that the latter warranted a higher score than the absence of

TABLE I

PER CENT OF LOWEST, MIDDLE AND HIGHEST THIRDS OF THE FAMILIES CLASSIFIED ACCORDING TO HOUSING SCORE *A* POSSESSING SPECIFIED HOUSING ITEMS*

Housing items and rank according to column 6	Per cent of families possessing specified items				Average "*t*" value of percentage differences between thirds
	All families	Lowest third "poorest"	Middle third "medium"	Highest third "best"	
1	2	3	4	5	6
1 Inside toilet with sewage system	46	4	42	94	11
2 Fully equipped bathroom	44	2	37	92	11
3 Electric sweeper	50	8	53	88	9
4 Central heating system	44	6	44	82	8
5 Running water in house	61	25	64	95	8
6 Monthly rental value of $20 or more	33	3	27	69	7
7 Sale value of house $1,500 or more	67	32	73	96	7
8 At least ½ closet per person	46	15	44	78	7
9 Kitchen sink with drain	78	48	88	99	6
10 Electric or gas refrigerator	41	11	42	70	6
11 Electric iron	82	55	93	99	6
12 Separate living room, dining room, kitchen and bedroom	62	33	65	87	6
13 Exterior paint unchecked	40	13	41	65	5
14 Electric lights in house	84	62	93	99	5
15 Telephone	27	5	18	48	5

* Items in this table are classified dichotomously and only the "desirable" category is shown. The table mentioned in footnote 5 shows a full list of categories of each item.

TABLE I—*Continued*

Housing items and rank according to column 6	Per cent of families possessing specified items				Average "*t*" value of percentage differences between thirds
	All families	Lowest third "poorest"	Middle third "medium"	Highest third "best"	
1	2	3	4	5	6
16 Walls of unbroken plaster	64	37	71	85	5
17 Walls—papered, painted or decorative plaster	73	47	79	92	5
18 Full basement with cement floor	21	4	16	43	5
19 Gas or electric cook stove	35	13	34	59	5
20 At least 1 room per person	88	70	95	100	5
21 Lawn and plantings in yard	63	37	67	83	5
22 Water system enclosed	70	48	72	91	5
23 Unbroken screens on windows	79	66	97	99	5
24 Not infested with vermin	64	42	65	84	4
25 House structurally sound	89	71	97	99	4
26 Roof of asphalt, asbestos, tile or slate	54	32	57	73	4
27 At least 200 sq. ft. per person	69	49	71	87	4
28 House of lapped planed siding, shingles, brick, stone, stucco, concrete or cinder blocks	90	73	98	99	4
29 Services within 1 mile	43	24	40	62	4
30 Steps of solid material	87	66	97	96	4
31 Blinds, drapes and/or curtains	91	75	97	99	4
32 Doors unbroken	90	75	97	99	4
33 Floors of grooved wood	87	74	89	98	4
34 Radio	91	77	98	99	4
35 At least ½ bedroom per person	88	75	91	98	3
36 At least 100 sq. ft. bedroom space per person	59	43	57	76	3
37 Power driven washing machine	80	61	85	93	3
38 House owned by family	64	49	61	81	3
39 Ceiling 8 to 10 feet high	83	67	89	93	3
40 No leaks in roof	85	71	90	94	3
41 Neighbors within ¼ mile	84	73	83	94	3
42 Foundation of solid walls	96	87	99	100	3
43 Woodwork painted or varnished	97	92	98	100	2
44 Unbroken glass in windows	95	86	99	99	2
45 At least 1 sleeping unit per person	93	86	95	98	2
46 House insulated	10	3	10	16	2
47 At least 20 sq. ft. of windows in living room	9	2	9	14	2
48 Kitchen cabinet with enclosed shelves	95	86	99	98	2
49 Power driven sewing machine	6	2	5	11	2
50 Floors covered	99	96	100	100	1
51 Garbage hauled or destroyed	98	94	100	99	1
52 House more than 100 ft. from barn, pens, railroad, factory, etc.	91	87	92	94	1
53 Hard-surfaced road adjoins lot	75	70	74	79	1
54 No obvious fire or traffic hazards	74	72	72	78	1
55 House under 10 years old	7	6	7	9	1
56 Single family structure	90	90	90	90	0
57 Working place within 2 miles	62	67	55	64	0

mechanical sweeping facilities; moreover, the author assigned the weights of 6, 3 and zero to these three categories (types) of the item, sweeping facilities. The assumption that sweeping facilities should be given approximately the same weight in the total as each of the other 56 component items was protected by limiting the range of scores from zero to 6, assigning the highest at 5 for dichotomous items and 6 for items of 3 · · · *n* categories and the lowest at 1 or zero. *A priori* weights were assigned to the various categories of each of 57 component items for which information was obtained. Table I indicates the 57 items. The total housing score for a family was the arithmetic average of the weights assigned to each item.

$$A = \frac{\sum a_k}{n},$$

where A represents the composite score, a_k represents the *a priori* weight assigned to item k, and n represents the number of items for which information was available (57 in most cases). In assigning these weights other studies together with the evaluations of other persons were drawn upon in the hope that the method, which involves much less work than other methods, would prove reliable when subjected to empirical verification by methods *B*, *C* and *D*. The *A* scores of the families studied ranged from 1.8 to 5.5.

Method B was designed to overcome one of the shortcomings of Method *A*; namely, that it does not provide differential weights for those desirable categories which are common in contrast to those which are rare. For example, 97 per cent of all dwelling units had painted or varnished woodwork (a desirable category of one of the 57 components) as compared with 6 per cent of the total that had electric sewing machines (a desirable category of another component). To assign weights empirically that quantitatively represented frequency of occurrence of categories (in terms of desirability), the percentages of dwellings in each category were converted into standard-deviation units of the normal probability integral. The procedure involved first the arraying of the categories of each component in order from least desirable to most desirable. For each of the 57 components the array of categories was exhaustive, that is, it covered 100 per cent of the categories. In the accumulated array, therefore, the least desirable category fell between zero and the percentage of dwellings in the category. The most desirable category fell between 100 per cent and 100 minus the percentage of dwellings in the category. Mid-points of these percentage ranges were then converted into standard-deviation units. The whole procedure is illustrated in an example. No mechanical sweeping facilities were present in 42 per cent of the houses, hand (non-powered) sweepers in 8 per cent, and electric sweepers in 50 per cent of the dwellings. The accumulated percentages were 42, 50 and 100; the mid-points (which were used to represent the category) were .21, .46 and .75. These mid-point values minus .50 (since area-tables are in terms of .50 of the area under the normal curve) were −.29, −.04 and +.25; the corresponding standard deviation units were −.8, −.1 and +.7; and the final weights (the sigma scores plus a constant 3.0 to eliminate negative signs) were 2.2, 2.9 and 3.7. The composite index for a dwelling was the arithmetic average of its scores for each component item.[2]

$$B = \frac{\sum b_k}{n}$$

where B represents the composite index, b_k represents the score on component k expressed in sigma units, and n represents the number of components for which information was available (57 in most cases).

Method C was designed to overcome a second criticism of scale *A*; namely, that it does not give greatest weight to those components that are most diagnostic of (most highly correlated with) the total score. The procedure involved the use of the total score obtained by Method *A* as a preliminary criterion upon which all families could be classified as having best, medium and poorest housing. It was then possible to determine which categories of each component were most and which were least common among the three classes of houses. In the category "inside toilets," for example, were 4 per cent of the poorest, 42 per cent of the medium and 94 per cent of the best classes of houses. To determine the significance of the differences among the successive classes, critical ratios were calculated for each of the 57 components between (a) poorest and medium and (b) medium and best. The final weight for each item was, then, the arithmetic average of the two critical ratios. The total score for each dwelling was the arithmetic average of the weights for the 57 items.[3] Table I shows the average "*t*" values of each item.

---

[2] For other examples of the application of this method see A. M. Leahy, *The Measurement of Home Environment*, Minneapolis: University of Minnesota Press, 1936; W. H. Sewell, *The Construction and Standardization of a Scale for the Measurement of the Socio-Economic Status of Oklahoma Farm Families*, Stillwater: Oklahoma Agricultural Experiment Station Technical Bulletin 9. 1940; H. R. Cottam, *Methods of Measuring Level of Living, Social Participation and Adjustment of Ohio Farm People*, Columbus, Department of Rural Economics and Rural Sociology, Mimeographed Bulletin 139, 1941.

[3] For other examples of the use of this method see Leahy, *op. cit.*, Sewell, *op. cit.*, and Cottam, *op. cit.*

$$C = \frac{\sum c_k}{n},$$

when C represents the composite index for a family, c_k represents the weight of the kth item, and n represents the number of components for which information was available. A serious disadvantage of this method is that each of the 57 components was forced into a dichotomous classification, whereas methods *A*, *B* and *D* permitted several categories and therefore greater refinement in weighting.

Method D was used to meet the chief criticism of Method *A*; namely, that the use of *a priori* weights is based upon subjective evaluations. A series of successive revisions of the *a priori* weights of each component were made in terms of averages of total scores of all 57 components for all dwellings. The method rests upon the belief that the usefulness of any single component item for predicting total housing can be repeatedly improved by using each successive revision of the total scores in revising the weights of the component items. Successive approximations were repeated until there was no marked change in the weights of the 57 component items from one revision to the next. More explicitly, the weights of Method *A* were temporarily accepted as first approximations which then required empirical revision to eliminate subjective evaluations. These weights of each category of each component were then revised by taking as the second-approximation weights the mean total housing score (M_N) of all dwellings reported in the Nth category of the component (i.e. all dwellings having the Nth category). These M_N scores were, however, first converted into comparable units with the mean score of all families M_D as a stable point of reference and with a $.3\sigma_D$ as a divisor which would give a reasonable range to the weights. Let O_N represent the weight to be assigned to the Nth component-category; let σ_D represent the standard deviation of the total housing scores; and let 5 be a constant. Then

$$O_N = \frac{M_N - M_D}{.3\sigma_D} + 5.$$

Now solving for M_N to simplify conversion to O_N by preparing a table of corresponding M_N and O_N values,

$$M_N = .3\sigma_D O_N - 1.5\sigma_D + M_D.$$

Finally, the total score (D) was

$$D = \frac{\sum O_N}{n},$$

where n represents the number of components for which information was available (57 in most cases). Consider the case of one component item, for example. The mean A score of all dwellings having an electric sweeper was 5.34, the mean score of those having hand (non-powered) sweepers was 5.08 and the mean of those having no mechanical sweeping facilities was 4.15. These were the M_N weights used in the second approximation, and upon the basis of them a new composite score D_2 was calculated ($D_1 = A$). Upon the basis of D_2 new weights were again assigned and D_3 was calculated. Similarly D_4 was calculated.[4] At the fourth approximation D_4, the weights had stabilized to a point where there was little change from one approximation to the next. And

$$r_{D_1D_2} = .983; \qquad r_{D_2D_3} = .995; \qquad r_{D_3D_4} = .996.$$

The principal advantage of this method is that items which bear no empirical relationship to housing are automatically reduced so that they exert no effect on the total score. Moreover, reliability and internal consistency are maximized. The main shortcoming is that categories which are infrequently present (i.e., low frequency) are subject to great chance fluctuations which are not corrected by the successive approximations.

Methods A, B, C, and D were found to result in strikingly similar housing scores for the sample families although the weights for various categories of the 57 items differed in some respects, particularly in those items which were not highly diagnostic of the total score. A table showing the weights assigned by methods A, B, C, D_1, D_2, D_3, and D_4 to each category of each of the 57 items has been prepared.[5] The intercorrelations (Pearsonian r) of the housing scores found by the four methods are as follows:

$$r_{AB} = .96 \qquad r_{BC} = .94$$

$$r_{AC} = .95 \qquad r_{BD_4} = .95$$

$$r_{AD_4} = .96 \qquad r_{CD_4} = .96$$

For purposes of constructing abbreviated scales of rural housing that might be standardized by subsequent testing, the weights adopted were those assigned by Method A with slight revisions to decrease the number of categories in certain items. The weights used in the abbreviated scales are shown in Table II. The criterion for selection of items to be

---

[4] For a more detailed statement of this method see Charles I. Mosier, *Measurement in the Field of Rural Housing* (typewritten) University of Florida. June 1941, pp. 32–37; 153–156b.

[5] Copies of this table in mimeographed form can be secured from the author.

TABLE II

WEIGHTS FOR ABBREVIATED HOUSING SCALES OF VARYING LENGTHS

(Combine first 10, first 20, or all 30)

Score $= \Sigma s/N$ where s is score of each item and N is number of items included

Rank of item	Housing item and weight
1	Toilet facilities: none (0); outside (2); inside—outside surface drain (3); inside—leechy cesspool (4); inside—septic tank(5); inside—public sewage (6).
2	Bathroom equipment; none (0); room, no fixtures (2); tub or shower only (4); lavatory only (4); shower or tub and lavatory (6).
3	Kind of sweeper: none or broom (0); hand carpet sweeper (3); electric sweeper (6).
4	Kind of heating system: fireplace, wood-coal (0); range (1); heater, wood-coal (2); gas heater (3); pipeless furnace (4); piped air (5); steam or water (6).
5	Water conveyance system: carry (0); hand pump, porch (1); hand pump, inside (2); running, cold (4); running, hot (6).
6	Monthly rental value: \$0–4 (0); \$5–9 (1); \$10–14 (2); \$15–29 (3); \$30–39 (4); \$40–49 (5); \$50 and over (6).
7	Sale value of house: \$0–249 (0); \$250–499 (1); \$500–1,499 (2); \$1,500–1,999 (3); \$2,000–3,999 (4); \$4,000–4,999 (5); \$5,000 and over (6).
8	Closet facilities per person: 0–.4 (0); .5–.9 (1); 1.0–1.4 (2); 1.5–1.9 (3); 2.0–2.4 (4); 2.5–2.9 (5); 3.0 and over (6).
9	Kind of kitchen sink: none (0); undrained (1); drained to surface (2); drained to sump (4); sewage system (6).
10	Refrigeration: none (0); spring house or cellar (2); ice (4); electric or gas (6).
11	Ironing facilities: none (0); flat (2); fuel (4); electric (6).
12	Room combination: separate living room, dining room, kitchen and bedroom (6); separate bedroom, dining room combined with living room or kitchen (4); other combinations (0).
13	Exterior paint: none (0); whitewash (2); cracked or checked paint (4); unchecked paint (6).
14	Lighting system: candles (0); kerosene (1); mantle lamps (2); gas (3); electricity (6).
15	Telephone in house: yes (5); no (1).
16	Inside wall material: uncovered (0); paper (1); unmatched wood (2); matched wood (3); wall board (5); broken plaster (4); unbroken plaster (6).
17	Wall decoration: none (0); whitewash (2); kalsomine (4); torn wall paper (3); untorn paper (5); paint or decorative plaster (6).
18	Basement: none (0); part, dirt (2); full, dirt (3); part, cement (5); full, cement (6).
19	Kind of cook stove: fireplace (0); makeshift stove (1); wood-coal range (3); oil or gasoline (4); gas (6); electricity (6); other (3).
20	Room per person in winter: 0–.4 (0); .5–.9 (1); 1.0–1.4 (2); 1.5–1.9 (3); 2.0–2.4 (4); 2.5–2.9 (5); 3.0 and over (6).
21	Lawn and plantings: lawns and plantings (6); lawns or plantings (3); neither lawns nor plantings (0).
22	Water source: open stream (0); dug well (2); drilled well (4); cistern or enclosed spring (5); public system (6).
23	Condition of window screens: none (0); broken (3); unbroken (6).
24	Structural soundness of house: sagging or rotting (1); solid (5).
25	Roof material: paper (0); sheet iron or tin (2); wood shingles (4); composition shingles (5); tile or slate (6).
26	Floor space per person: 0–99 feet (0); 100–199 feet (1); 200–299 feet (2); 300–399 feet (3); 400–499 feet (4); 500–599 feet (5); 600 feet and over (6).
27	Exterior wall material: log (0); battened or tar-paper-covered boards (2); lapped-planed siding or shingles (4); stucco (5); concrete, stone, brick or cinder blocks (6).
28	Distance to services: under 1 mile (6); 1.0–1.9 miles (5); 2.0–2.9 miles (4); 3.0–3.9 miles (3); 4.0–4.9 miles (2); 5.0–5.9 miles (1); 6.0 miles or more (0).
29	Condition of steps: none (0); broken boards (1); loose or insecure materials (3); solid materials (6).
30	Window coverings: none (0); curtains or blinds (3); blinds and drapes and/or curtains (6).

included in the abbreviated scales was the diagnostic ability of each item as determined by Method *C*. More specifically, the 10 items used in the shortest scale (the first 10 in Table II) are the 10 which had highest average "*t*" values (Table I). To make the scale of 20 items the highest 20 items in Table I were used. Similarly, the first 30 items comprised the longest of the three abbreviated scales, except that the item ranking 31 was substituted for item 24 which was a difficult item for which to obtain information. The score for a house is the arithmetic average of the 10, 20 or 30 items, depending upon which scale is being used.

The *validity* of the abbreviated scales is indicated in the following facts: (a) Pearsonian correlation coefficients of each abbreviated scale and the longer scale of 57 items (Method *A*) were as follows:

$$r_{57.10} = .89; \qquad r_{57.20} = .94; \qquad r_{57.30} = .97$$

(b) The validity of the weights used in Method *A* (the source of the weights in the abbreviated scales) is indicated in the high intercorrelations of methods *A*, *B*, *C* and *D*.[6]

The *reliability* of the abbreviated scales is indicated in the following facts: (a) high internal consistency of the component items since they were selected as the most diagnostic items according to Method *C* and since all of the 57 items from which the weights were made by Method *A* showed high internal consistency. More specifically, a split-half test of the 57 items as weighted by Method *A* resulted in $r = .85 \pm .14$ and $\bar{r}$ (Spearman-Brown) $= .92$, and intercorrelation coefficients of scale *A* and sub-indices[7] of sanitation (*G*), safety (*H*), labor-saving devices (*I*), electrical equipment (*J*), and space (*K*) were as follows:

$$r_{AG} = .90 \pm .04; \qquad r_{GH} = .75 \pm .04; \qquad r_{HI} = .66 \pm .04$$
$$r_{AH} = .79 \pm .04; \qquad r_{GI} = .94 \pm .04; \qquad r_{HJ} = .61 \pm .04$$
$$r_{AI} = .86 \pm .04; \qquad r_{GJ} = .83 \pm .04; \qquad r_{HK} = .31 \pm .04$$
$$r_{AJ} = .77 \pm .04; \qquad r_{GK} = .34 \pm .04; \qquad r_{IJ} = .89 \pm .04$$
$$r_{AK} = .52 \pm .04; \qquad r_{JK} = .16 \pm .04; \qquad r_{IK} = .26 \pm .04$$

---

[6] Another test of the validity of scale *A* was made by calculating a coefficient of correlation between the scores of that scale and scores obtained by ratings made from photographs of the same houses in which 20 selected raters sorted 70 photographs of houses into seven piles representing temporary scores ranging from zero to six with the mean value of the 20 ratings becoming the final photo score of each house. In these ratings a test of reliability of the raters was made by including two photographs of seven of the houses (viewed from different perspectives). The ratings of those raters who proved to be reliable were used in determination of the final photo scores. The coefficient of correlation (Pearsonian *r*) of Scale *A* and the photo scores was $r = .92$ with a sampling error of .16 ($E_{517} + E_{70}$).

[7] There was some overlapping of the components included in sub-indices *G*, *H*, *I*, *J* and no correction was made for the spuriousness thereby introduced. The amount of spuriousness is, however, very

Also the mean housing scores on scale *A* for 20 matched families interviewed by two different interviewers were not significantly different.

$$M_1 = 4.23; \qquad M_2 = 4.25$$

(b) Reliability coefficients of the abbreviated scores by split-half tests were as follows (sampling error was .14):

$$r_{10a.10b} = .82; \qquad r_{20a.20b} = .86; \qquad r_{30a.30b} = .84$$

which when corrected by Spearman-Brown formula were:

$$\bar{r}_{10a.10b} = .90; \qquad \bar{r}_{20a.20b} = .92; \qquad \bar{r}_{30a.30b} = .91$$

(c) Questionnaires which were administered to 1,641 school children in the communities in which the interviews were made indicated that the interview data were recorded with reasonable accuracy.

To expedite comparison of subsequent data that may be obtained by use of the abbreviated scales with the scores obtained in the present study, decile and quartile values of abbreviated scales of varying lengths are shown in Table III.

TABLE III
DECILE AND QUARTILE VALUES OF ABBREVIATED HOUSING SCALES

Decile and quartile	Abbreviated scale		
	10 items	20 items	30 items
First decile	1.0	2.2	2.4
Second decile	1.3	2.5	3.0
First quartile	1.5	2.7	3.2
Third decile	1.8	2.9	3.3
Fourth decile	2.3	3.2	3.7
Fifth decile (Q_2, Md.)	3.0	3.6	4.0
Sixth decile	3.8	4.1	4.4
Seventh decile	4.3	4.5	4.7
Third quartile	4.5	4.7	4.8
Eighth decile	4.7	5.0	4.9
Ninth decile	5.1	5.3	5.2

In conclusion, it appears that the above abbreviated scales comprised of 10, 20 or 30 items may be used to measure housing in rural Pennsylvania with reasonable validity and reliability. Since the selection of items for the abbreviated scales was based on diagnostic ability of the items, very little validity and reliability are sacrificed by reduc-

---

small. *G* included items 1, 2, 3, 4, 5, 9, 10, 11, 14, 18, 19, 22, 23, 24, 37, 40, 44, 45, 47, 48, 50, 51, 52, 54; *H* included items 4, 14, 16, 22, 23, 25, 26, 28, 30, 33, 40, 42, 44, 52, 53, 54, 55; *I* included items 1, 2, 3, 4, 5, 9, 10, 11, 14, 15, 19, 37, 48, 49; *J* included items 3, 10, 11, 14, 15, 19, 34, 37, 49; and *K* included items 8, 12, 20, 27, 35, 36, 45.

ing the number of items to 10. Standardization of the scales awaits further testing.

The research described in this article suggests possibilities of constructing scales for other populations. Similar work could be done with secondary data from the *1940 Census of Housing*. Another statistical problem that, according to the above analysis, may deserve further attention is the comparative usefulness of *a priori* weights in constructing composite indices and weights assigned by other means.

COORDINATING PRODUCTION FOR WAR*

BY SHELBY CULLOM DAVIS

THE CLASSIC NIGHTMARE at our auto plants used to be the spectacle of a four-door body arriving at the final assembly line just in time to be lowered on a chassis built for two. The nightmare was the worse if the body was red and the chassis green. But such spectres practically never occurred. The reason: coordinated production. Every part, every component in the right place at the right time—a constant flow of materials through the plant and resulting economies in the use of materials, warehousing space, and dollars to finance inventory. Coordination in the manufacturing process meant economy of facilities, too. All departments had a relatively even capacity and some were not idle, further tying up investment and making for unbalance.

What has been urgently needed in the American war economy is just such coordinated production. During the spring and summer of 1943 it has been slowly emerging—a new phase in the war production effort of World War II, an immensely valuable phase whose rewards are measured in the thousands of additional airplanes, guns, tanks, ships and other items which hasten the day of ultimate victory.

Statisticians, both in Washington and in the field, have played an invaluable part in bringing about this coordinated production and ensuring its success through insistence upon accurate and realistic production schedules. Although it is still too early to point, chapter and verse, to the statisticians who fought through the "Battle of Washington" to achieve these ends, no one can deny the contribution which statistics has made in forecasting the bottlenecks and freeing the economy from the chaos which frequently threatened.

Three years ago, in the days of the National Defense Advisory Commission which was set up after the fall of France in 1940, the great bottleneck was machine tools. The term "bottleneck" became an everyday part of the vernacular. Now most of these machine tools are in place, and are hungrily devouring all kinds of materials. Other bottlenecks, however, threatened to become so numerous that complete coordination of production, as in the automobile plants, became imperative. That the war production cycle is fast moving toward its final phase, indeed, is indicated by the fact that machine tool backlogs have for some time been declining and many machine tool makers are starkly facing the problem of conversion to other war work.

---

* A revised version, brought up to date, of a paper presented by Mr. Davis at a meeting of the Industrial Statistics Group of the New York Chapter on February 17, 1943. The author at that time was Chief Statistician of Region II (New York State—Northern New Jersey) of the War Production Board; now he occupies a similar position for the Smaller War Plants Corporation.

Coordination of production is the third and final stage of our war economy. The first stage was the make-ready stage. The emphasis was upon more and ever more machine tools and the construction of new plants, such as the Chrysler Tank Arsenal. Our arsenal for democracy was superimposed upon the structure of an already booming civilian economy.

The second phase—conversion—which occurred shortly after Pearl Harbor, was dramatized by the shutdown of the automobile industry in February 1942. L and M Orders limited the production of goods to a certain level, or eliminated them entirely. The main purpose of the economy became war. Contracts were fairly shoveled out by the government procurement agencies. The production curve climbed steadily. Then it began to falter. The answer was, in part, materials shortages and the problems of scheduling as these shortages increased. More steel, copper and aluminum were being produced than ever before, but not enough.

So the next big job was brought to focus: achieving maximum sustained production by helping to get every plant on an approved schedule within a precisely balanced over-all program. Materials and other resources had to be budgeted and set aside and allocated for the specific purpose of producing the greatest possible amount of the things needed most by the United Nations at any given time.

The normal laws of supply and demand had long since become inoperative, because there was no ceiling to demand. The Nation's ability to produce, which had never before been tested, was definitely limited. The preference rating system, designed to insure that urgent production got materials ahead of less urgent, bogged down. Priority certificates, which authorized their holders to receive equipment or materials, became known as mere "hunting licenses"—for frequently more certificates were outstanding than the equipment or material to honor them.

The Production Requirements Plan, a big improvement, was set up in the late spring of 1942 as a means of allocating materials on the basis of the quarterly needs of the various manufacturers. It proved inadequate principally because it did not compel the total program to be kept within ability to produce. And it didn't make sure that some components wouldn't be manufactured at the expense of others. The patriotic manufacturer, for example, who boasted his production of machine guns was 60 per cent ahead of schedule, had unwittingly caused some other manufacturer of an equally important product to be behind schedule. Further, some manufacturers, determined to keep their plants on schedule, ordered more materials than they needed before they needed them. The temptation to over-order, in general, proved too strong in too many instances. As a result, materials piled up in inventory, instead of flowing swiftly through the country's machine shops and factories.

Meanwhile, individual delivery schedules were being set on all prime contracts awarded by the Army, Navy or Maritime Commission. When the Navy ordered a warship, it specified the completion date and a time schedule leading up to fulfillment of that completion date. The same for tanks, machine guns or airplanes. But no over-all authority existed to supervise all scheduling. No one saw, first of all, that the scheduling was possible in terms of materials and facilities available, and secondly, that the scheduling was balanced from the point of view of military necessities. Furthermore, schedules of production more often than not were too optimistic and represented the wishes of the procurement agencies rather than an accurate, statistical measure of their fulfillment. And since the procurement agencies awarded all the contracts, they made up all the production schedules on prime contracts. Schedules on subcontracts were, of course, made by the individual prime contractors, but they were usually too optimistic also and frequently called for deliveries long before they were actually needed.

The Army-Navy Munitions Board did attempt to act as a kind of umpire between the two greatest sources of materials demand, the Army and Navy, but its efforts were only partially successful. Clearly a third party, representing the country as a whole and acting as an objective arbiter, was essential.

By the autumn of 1942, it had become apparent that production throughout most of 1943 would lack a tight and well developed system of distributing materials. The Production Requirements Plan would end in the first quarter, and the new Controlled Materials Plan (CMP), which aimed at tightening control of materials through allocating the exact amounts of steel, copper and aluminum needed for end products, would be introduced in the second, but it would probably not become effective until the third. It was felt that the coming months would witness a general excess of fabricating facilities while raw materials would act as the general limiting factor. Nevertheless, many of those in authority believed that important fabricating bottlenecks retarding production of components and end items would arise. Hence, such bottleneck items as compressors, valves, bearings, generators, Diesel engines and certain machine tools should be given special attention. Furthermore, while the schedules were being brought in line with fabricating capacity and raw materials, they should be reviewed concurrently in relation to manpower resources.

Although production had advanced steadily throughout 1942, there was increasing evidence of widespread imbalance in the production program. Production of many end products had run far ahead of schedule, while for others it had fallen far behind. Among complementary items such as guns and radios in relation to tanks and airplanes, and ammunition in relation to weapons a similar situation prevailed. So long as an adequate supply of raw materials and other resources existed, this imbalanced production was not serious. But with machine tools, and labor growing scarcer, needless hoarding of raw materials in inventories of components or end products seriously aggravated the situation.

In the case of aircraft, for example, schedules were optimistic, partly to set high goals as an incentive. Certain branches of the Services presented forecasts of production which were not attained by as much as 50 per cent month after month. Although it might have been desirable to keep the schedules high in order to serve as incentives, the excess of forecasts over production had been too great to achieve this end. Ground Army equipment production was understood to be falling increasingly behind the schedules, particularly such difficult items as self-propelled artillery. Students of the problem thought that each manufacturer or procurement agency should schedule out production of end items, parts and components on the basis of realistic production probabilities of industrial plants in existence or being constructed.

As the new Controlled Materials Plan, previously mentioned, was being born, under the guidance of Ferdinand Eberstadt, WPB Vice-Chairman, it seemed obvious that the flow of materials could not be efficient and successful if the schedules which had already proven so erroneous were the basis for the material requests of the Claimant Agencies. Likewise, strategic decisions certainly could not be based upon the given schedules which past performance had revealed to be so substantially in error.

On December 8, 1942, the WPB Industry Divisions were asked by Mr. Ernest Kanzler, at the request of the Military Committee under the chairmanship of General Lucius Clay, to get from Claimant Agencies their requirements for common components. Only Rubber and one other agency were able to reply, and they only in part. If the Claimant Agencies, such as Army, Navy, Maritime, or Lend-Lease, did not have a thorough grasp of these vital facts, who would? Here was proof positive that a gaping hole of ignorance had been opened up which badly needed to be plugged by realistic statistics and coordinated production.

On December 31, Mr. Ralph Cordiner, whose Scheduling Committee replaced this Military Committee, asked the Divisions concerned to call in the manufacturers of the critical components and appraise the situation. This developed (1) a shortage of materials and (2) a shortage of orders. With Mr. Eberstadt's cooperation, the necessary materials were arranged for.

The order shortages were more difficult. One industry reported on January 11, 1943, that its unfilled orders for the second quarter of 1943 were only 40 per cent of its production in the fourth quarter of 1942. Suppliers of critical components, facing declining backlogs of unfilled orders, were reluctant to step up current production rates still further, even if requirement estimates of the Services showed needs taxing capacity for the year ahead. During January an attempt was made to have orders placed at once for critical components, before February 6 for delivery to June 30 and before March 1 for delivery in 1943—but the furor which this circular letter of January 28 to several thousand builders created, showed the need. In the case of one prominent manufacturer of ammunition components, firm orders extended only through February.

Donald M. Nelson, Chairman of the War Production Board, was thoroughly aware of the growing need for scheduling and of the fact that actual war production in most categories fell short of scheduled production throughout most of 1942, and increasingly so as the year advanced. In a report to President Roosevelt, he cited the fact that during the first quarter of the year actual production of two-engine medium bombers was 137 per cent of the forecast for the first quarter made on January 1; the corresponding percentage for the second quarter was 95; for the third 84; and for the fourth 72. Actual production, however, had increased almost constantly. The same was true of many other items.

There were several explanations for this falling behind schedule, according to Mr. Nelson:

1. The objectives were too high.
2. As critical resources became scarcer, schedules were boosted in order to strengthen the claim of the competing services on such resources.
3. The coordination of the country's resources in an attempt to meet these schedules was insufficiently rigorous.

4. The schedules were not subjected to a centralized review and control. This decentralization existed not only among, but also within, agencies.

Had schedules been prepared realistically, it would have been obvious early in 1942 that the objectives were out of line with resources and the goals could have been brought down to the limits of feasibility, thereby requiring early in the year a complete review of the entire program. Schedules were "seriously defective" as analytically useful predictions of war output, concluded Mr. Nelson.

Scientific scheduling actually began when the Production Scheduling Bureau of WPB examined the outstanding orders for common critical components in February and March, 1943, and with the aid of statisticians, began to schedule their production. In many cases this involved spreading the work. So-called task forces of industrialists, consisting of advisory scheduling committees attached to WPB industry divisions, made recommendations on transfers of orders for critical common components from firms having too large a backlog to others without enough work to go around on a continuous 1943 production schedule.

Attention was concentrated upon production of gears, valves, Diesel and gasoline engines, crankshafts, compressors, pumps, heat exchangers, welding rods and electrodes, electric motors, starters and generators, boilers, vacuum tubes and control instruments.

Many examples of thoroughly unscientific scheduling were found. One large builder of direct current motors, equipped to produce about 220 per month, was being called upon to ship more than 1500 within less than 30 days. At the same time a smaller manufacturer, with a capacity of more than 100, was being called upon to turn out only 60 a month.

The General Scheduling Order, M-293, issued in March, required all manufacturers of 47 major groups of components to submit monthly operating reports covering backlogs, new orders, deliveries, cancellations. M-293 provided the means of closing the gap between the Controlled Materials Plan, covering allocation of raw materials to manufacturers, and the delivery on schedule of the finished munitions. In effect, it did for critical common components what CMP did for raw materials—insured their delivery at the right place at the right time.

At a staff meeting shortly after WPB had assumed responsibility for over-all scheduling, Charles E. Wilson, Executive Vice Chairman of WPB, stated that a well-coordinated war production program would increase over-all production by something like 10 per cent; but that it could increase production of critical items, such as the five then current "musts" (escort vessels, 100 per cent octane gasoline, synthetic rubber, airplanes, and merchant ships) by as much as 25 to 30 per cent.

It is difficult, if not impossible, to know how much over-all gain in production has been achieved. Too many other factors, such as the normal production increase previously scheduled for 1943 or the growing tightness of labor supply, have intervened. It is correct to state, however, that many individual bottlenecks have been eliminated and threatened breakdowns in production averted. A three months' delay in the opening of a lens-making plant was avoided by persuading the management to accept delivery of compressors from a smaller, less widely known company. The delivery of heat exchangers for a large synthetic rubber plant was pushed forward three months by transferring $350,000 of orders from an overloaded producer to one operating below capacity.

Individual manufacturers are enthusiastic about the scheduling program, which they feel has brought about at least a reasonable semblance of order from chaos. A small company in northern New Jersey is a typical example. Before scheduling went into effect, this producer of electric motors (both alternating and direct current, from fractional to 200 h.p.) was frequently able to ship monthly only as much as one-third of what was desired. Furthermore, it had been common practice of the Claimant Agency and the prime contractor and, if one existed, the first tier subcontractor, each to advance the date on which the electric motors were required from the company by several months. Thus, company officials often suspected—and it was all too frequently true—that some of the motors on which it was lavishing herculean production efforts might repose in inventory at the point of final assembly of the finished munitions product for as much as six months. On the other hand, tardiness in delivery on another order might actually impede the war program.

With heavy but attainable schedules now in effect, the company has in turn been able to install an efficient internal scheduling system which insures that its monthly schedules are closely adhered to. In June, the first month of scientific scheduling, the company shipped 92 per cent of its commitments, a tremendous increase over its previous performance. Wide publicity was given in the plant to the production schedules for the twelve various types of motors. A large sign indicated, where everyone could see, the scheduled production for each type of motor, the shipments each week, and the number of units "to go" during the remainder of the month.

The officials of this relatively small plant predict scheduling will still further increase production by 25 per cent within a few months. Other examples, involving larger companies, could also be cited.

With production of some critical munition items scheduled to increase as much as 50 per cent during 1944, the need for the most efficient operation possible—as if America were one gigantic automobile plant, where scheduling has long since been reduced to a science—is obvious. Only through the most rigorous scheduling, aided and abetted by the Nation's statisticians, will America be geared to produce just what is needed, when it is needed—not too little and too late—but enough, and on time.

STATISTICS FOR WAGE STABILIZATION

By Robert J. Myers and Harry Ober
U. S. Bureau of Labor Statistics

The administration of the Federal program for the stabilization of wages has required the accumulation of an unprecedented volume of detailed and systematic wage information. The primary source of such information is the Occupational Wage Rate Project, originated early in 1943 by the Bureau of Labor Statistics[1] and financed in part by the National War Labor Board. This project not only serves as a useful instrument of wartime economic control, but also provides a type of wage statistics long needed in collective bargaining.

The wage statistics generally available from public sources are designed primarily for other purposes and are too broad in scope and too general in nature for effective use in wage stabilization or the settlement of wage disputes. For these purposes, wage statistics must be available by individual occupation and by particular locality. The collection of such data is laborious and expensive, and in normal times has been considered beyond the means of public agencies. The lack of impartial statistics on occupational wages has seriously impeded the peaceful negotiation of wage disputes.

On several occasions in the past the Federal Government has undertaken to provide occupational wage statistics for special purposes. Unfortunately, the usefulness of these statistics has been considerably reduced due to delay in determining their need, and to the length of time required to plan and accomplish their collection. Thus, the most ambitious study of occupational wages in World War I, undertaken at the request of the War Industries Board, was not started until the summer of 1918 and did not result in a preliminary report until the spring of 1920. Special wage studies undertaken for the National Recovery Administration, although yielding early and useful results, again demonstrated that detailed and accurate wage statistics cannot be produced overnight.

The outbreak of World War II found comparatively little information available from public[2] sources regarding wages in specific occupations and localities. General industry surveys by the Bureau of Labor Statistics provided information on earnings by occupation in a few industries, but these studies permitted analysis only by broad regions. Common-labor entrance rates were known for the largest cities, as were union wage rates for a few important jobs. A small amount of additional information was available in the files of various administrative agencies. For most industries, however, and particularly the nonmanufacturing industries, up-to-date occupational wage rates by individual city were largely or entirely lacking.[3]

---

[1] The Occupational Wage Rate Project herein described was developed under the general supervision of A. F. Hinrichs, Acting Commissioner of Labor Statistics, and N. Arnold Tolles, Chief of the Bureau's Working Conditions and Industrial Relations Branch. The project is operated by the Wage Analysis Division of the Branch. Robert J. Myers is Chief of this Division and Harry Ober is immediately in charge of the project. Lily Mary David, assistant supervisor of the project, has contributed substantially to its planning, while Toivo Kanninen and John F. Laciskey have been responsible for the preparation of the large number of required job patterns.

The shortage of occupational wage data was first acutely felt in 1941 by the National Defense Mediation Board, which was responsible for the settlement of wage disputes but not for general wage stabilization. This Board found it necessary to initiate a number of brief, special studies for the purpose of comparing wage rates in disputant plants with those in comparable plants in the same industry and/or community. These were "tailor-made" studies, designed to fit the particular case involved and, in some instances, covering various other items in addition to wages. A number of special studies were also requested by the National War Labor Board, which succeeded the N.D.M.B. in January 1942. Much of the special research was done by the Bureau of Labor Statistics, which undertook approximately 75 studies for the Board during the first nine months of 1942.

With the launching of the stabilization program in October 1942, the number of special studies multiplied, many of them being used in connection with the review of voluntary wage-change proposals. In the first quarter of 1943, in fact, the Bureau completed more than 700 separate studies, although the majority of these involved little or no field investigation. Meantime the number of applications for approval of wage changes rose above 10,000 per month, and it became apparent that a larger, more systematic, and more economical organization for gathering wage data was required. In anticipation of the need for the systematic collection of occupational wage rates, plans for such a study were made by the Bureau by the end of January 1943, and the project was launched in full scale early in April. Within the next six months the Occupational Wage Rate Project collected wage data from more than 60,000 manufacturing and nonmanufacturing establishments and transmitted summary reports covering more than 50,000 of these to the War Labor Board.

---

[2] Valuable material on wages in a few industries, particularly the metal-working industries, was obtainable from trade associations, individual companies, unions, and other private agencies. Some of this material had been collected from association members or other limited groups, and was not claimed to represent other segments of the industry. Material presented by interested parties, moreover, is often unacceptable to one of the participants in a wage dispute. Nevertheless, the wage material provided from private sources has served many purposes and has contributed substantially to the advancement of the wage-stabilization program.

[3] Apparently most other industrial nations were also lacking in occupational wage statistics. Canada, however, preceded the United States in the collection of such material. Occupational wage rates are available for important jobs in many Canadian cities and have been of considerable importance in administering the Canadian program for wage stabilization.

ESSENTIAL FEATURES OF OCCUPATIONAL WAGE STATISTICS

The characteristics of wage statistics for use in wage stabilization and in collective bargaining must, of course, be conditioned by the assumptions and policies of the administrative agency or the negotiating parties. It is of fundamental importance that wage determination in the United States is greatly influenced by common wage practice throughout the economy. Information regarding the prevailing level of wages is generally more useful than calculations of the cost of adequate budgets or similar measures. It is also highly significant that wage stabilization in this country has aimed at stabilizing wage rates, rather than weekly, monthly, or annual earnings.

Important, too, in its influence on wage statistics, is the unit (locality or industry) within which stabilization is to be achieved.[4] The importance attached to traditional wage differentials—for example, between jobbing and captive plants, or by sex or race of worker—is another factor that must be considered. It may be assumed, however, that policies of this type may change from time to time. Insofar as possible, the system of wage statistics should be flexible enough to accommodate itself to such changes.

In addition to those characteristics that reflect the particular policies of the administrative body, several important features would appear to be essential in almost any statistical system designed to determine prevailing wage rates for purposes of wage stabilization. First, the wage levels should be measured in terms of a stable and definite unit. Second, the data must be presented in extensive detail in order to reflect inter-area and inter-industry differences. Third, the wage statistics provided should be reasonably up-to-date. Fourth, in the interest of speed and economy, the statistical system should be simplified to the maximum extent consistent with the requirements of the administrative program. The methodological implications of these features are rather broad and justify brief discussion.

---

[4] In this respect public policy appears to have been considerably influenced by traditional practice. In shipbuilding, railroads, and certain other industries, local differences in general wage levels have been largely ignored. In the determination of wages in many industries, however, the locality approach has been dominant.

Unit of Measurement.—Statistics for wage stabilization and collective bargaining must permit frequent and direct comparisons of wages among groups of establishments, often across industry lines. It is essential, therefore, that a common unit of measurement be used. The most suitable unit is the hourly wage rate—or its equivalent, straight-time hourly earnings.

Weekly or monthly earnings are usually available, but these measures have been greatly influenced by variations in the number of hours of work, and hence constitute poor measures of the rate of remuneration. Even in nonmanufacturing industries, where salaries are commonly calculated by the week or half-month, differences in pay may reflect nothing more than differences in hours of work. Manufacturing establishments usually pay their workers by the hour or, in the case of "incentive workers," more or less in proportion to their production.

Incentive wage payment, whether consisting of straight piece rates or of a guaranteed base rate plus a bonus for output above a set norm, presents troublesome problems. Because of variations in the unit of output and in the characteristics of the different incentive plans, comparisons of the rates "per piece" are out of the question. For incentive workers, therefore, it is necessary to compute average hourly earnings over a period of time, and to substitute this average for the hourly rate obtained for time workers. For a variety of reasons, including the greater intensity of their work, incentive workers typically earn more per hour than time workers performing the same operations. Wherever practicable, therefore, the earnings of incentive workers should be tabulated separately from the rates of time workers.

Premium payments for overtime or late-shift work should not be permitted to distort wage-rate comparisons. Such payments should be considered as remuneration for special exertion or discomfort and dealt with separately from the rate problem. Christmas bonuses, vacation pay and similar items should also be considered separately and not combined with basic wage rates.

Wage rates are, of course, of little significance except in relation to the type of work performed. It is essential, therefore, that wage rates be collected and presented *by occupation.*[5] Some occupational groups, in fact, are so broad as to make further subdivision desirable.

Industry and Locality Boundaries.—Even for approximately the same type of work, wage rates differ substantially from industry to industry and from locality to locality. The causes of such differences are complex and do not permit discussion here. The effect on wages is substantial, however, and warrants the use of fairly narrow boundaries in ascertaining levels of wages.

---

[5] Weighted averages of the rates or earnings of all workers in a plant are sometimes compared during wage negotiations with similar averages for other plants. This device may be extremely misleading, however, because of differences in the "weight" of the various occupations. A plant which pays a higher rate than another in each separate occupation may show a lower plant average because of a difference in organization which permits the employment of relatively more unskilled workers.

Some occupational classifications appear in essentially the same form in several industries, and workers in these jobs may transfer readily from one industry to another. In such occupations inter-industry differentials may be minor,[6] and the most important boundaries to observe in determining wage levels will probably be geographical. On the other hand, approximately the same levels of wages may prevail in neighboring cities over an extended period of time. In general, however, both geographical and industry limits are important considerations in wage statistics.

Even when the occupation, the industry, and the locality are the same, significant wage differences may be found because of other factors, such as differences in size of establishment, union status, method of wage payment, and so forth. The number of establishments in a given industry and locality rarely permits segregation of more than one or two of these factors in the same tabulation.

Timeliness.—It goes without saying that statistics for use in wage stabilization or wage negotiations should be relatively up-to-date. This means that unnecessary complications in the collection or tabulation of the wage data should be avoided, since these may delay the progress of the work.

In times of considerable wage movement, frequent repetition of wage studies will be essential. In addition to providing an up-to-date picture of prevailing wages, repetitive studies of occupational rates permit the construction of *indexes of wage rates,* which measure fairly accurately the effectiveness of the stabilization program.

Characteristic Industries and Key Jobs.—A complete occupational analysis of wages in all industries and in all communities would obviously involve great delay and enormous expense. The Dictionary of Occupational Titles, issued by the U. S. Employment Service, lists almost 20,000 different jobs, most of which are found in hundreds of communities. Fortunately, wage statistics for a small proportion of these jobs and in relatively few communities can provide an adequate statistical basis for collective bargaining and for official wage determination.

---

[6] Inter-industry wage differences may be substantial, however, even for nearly identical types of work in the same locality. Such differences often reflect variations in the regularity of employment, the safety or agreeableness of working conditions, and so forth. Such factors undoubtedly account in part for the high wages in shipbuilding and the construction industry as compared with the wages of more or less comparable workers in manufacturing.

In the first place, the wage levels of many communities are greatly influenced by wages in a small group of *characteristic industries*, which, with a few additions, provide a reliable pattern of wage structure. The characteristic industries vary, of course, by community. Opportunities for inter-industry transfers, supported by local tradition, help to maintain the relationship between wages in the minor trades and those in the leading industries. In many communities a survey of only 10 or 12 industries will cover more than half of all employed workers, while wages in the remaining industries may be closely tied to those that have been studied. Moreover, wage observations for the remaining industries may apply to too few workers to justify systematic analysis. Brief special studies will, of course, be necessary in some cases to determine existing relationships.

The study of all jobs in a given industry is even less justifiable than the study of all industries. The relationship of wage rates for the various occupations within a given industry shows considerable uniformity and, even when modified because of special circumstances in a particular establishment, may be almost as stable as the organization of production itself. The disturbance of occupational differentials is in most establishments a serious matter, which will not be undertaken lightly either by management or the union. Hence, the determination of wages in a few *key jobs*, although somewhat less enlightening than a complete occupational analysis of the industry, provides a sound basis for wage adjustment. Unquestionably the study of occupational wage rates in terms of key jobs facilitates analysis and permits great financial economies.

The key-job approach is supported by successful experience in collective bargaining and in wage determination. Many unions and employers have negotiated wages in this manner for years, adjusting wages in the intervening occupations automatically once the rates for key jobs are agreed upon. Job-evaluation experts depend heavily on key jobs in rationalizing company wage structure. Early experience under the Federal wage-stabilization program indicates that the key-job approach is not only suitable for purposes of wage statistics but may also serve successfully as a method of wage control.

It is obvious that other economies may also be realized without seriously reducing the value of the system of wage statistics. The study of prevailing wage rates in certain industries in large cities may safely be based on sample establishments. Many small communities may be omitted completely if other communities in the same region have been covered.

CHARACTERISTICS OF OCCUPATIONAL WAGE RATE PROJECT

The practicality of the features discussed above has been demonstrated by the experience of the Occupational Wage Rate Project. As a product of the war situation, this project was designed primarily to meet the needs of the War Labor Board. Both the plan and operation of the project have been modified by the day-to-day demands of the stabilization program. Yet uniformity, particularization, timeliness, and simplicity have remained among its major characteristics.

Data Collected.—The basic wage information obtained through the Occupational Wage Rate Project consists of hourly wage rates or (in the case of incentive workers) straight-time hourly earnings of experienced workers in specific occupations.[7] In order to eliminate the influence of late-shift premiums, the study of rates and earnings is usually confined to first-shift workers, although the total number of workers in the occupation is used for weighting purposes. The occupations studied are defined in writing, with separate descriptions in some cases for different grades of work within the job. Thus assemblers and many other workers are identified as grade A, grade B, or grade C. All wage figures are obtained directly from pay rolls or other basic records, and represent a typical pay-roll period (usually a week, two weeks, or a half-month) during the months of January, April, July, or October. All wage information intended for combination is, of course, obtained as of the same period.

In addition to the basic wage data, a considerable amount of supplementary information is secured, including type of product, total employment, provisions for overtime, late-shift work and vacations, the existence of union agreements, and other items. This material is of great value in classifying and interpreting the wage data.

The proper classification of workers by occupation and grade is, of course, of primary importance. Unfortunately, the classification systems of individual companies differ widely, and frequently forbid direct comparison. The designation of occupation and grade is consequently made independently of local practice and requires the careful attention of a trained field representative. Such field representatives discuss the duties of each group of workers with the labor manager, foreman, or other informed person, and frequently observe the actual performance of the work. The training of new field representatives has constituted one of the most difficult problems connected with the Occupational Wage Rate Project. Its accomplishment has required the cooperation of a number of trade schools, much special office instruction, and experienced supervision on the job. At one time more than 325 field workers were engaged in the collection of occupational wage rates for this project.

---

[7] Specimen copies of the field schedules used (forms OWR-1 and OWR-2) will be supplied on request to the Bureau of Labor Statistics.

The tabulations produced by the Occupational Wage Rate Project for use by the War Labor Board,[8] a sample of which is reproduced herewith, are relatively simple in nature, but present substantial detail. For each key job in a covered industry and wage area they present the weighted average hourly wage rate and the range of rates for every establishment. Individual establishments are not identified, since cooperation in the project is on a voluntary basis and most companies are willing to report their wage rates only for use without identification. Insofar as is possible without revealing individual operations, however, the individual establishments are classified by size, unionization, and method of wage payment. Brief interpretative comments are transmitted with all tabulations, and special analyses are made on request or when obviously appropriate.

Scope.—As of September 1943, wage data have been collected from more than 400 separate localities, including virtually all cities of 25,000 population or more and a substantial number of smaller places. No attempt has been made to confine the collection of data within city limits, since similar wage levels often prevail in a large city and its environs, or in two or more neighboring communities. Places of less than 25,000 population have been covered largely on the basis of special administrative need, and with little regard for sampling theory.

In each wage area the intention has been to cover the characteristic industries. This plan has been adapted to the needs of the regional War Labor Boards, however, and as a result the representation of the characteristic industries has varied considerably from one wage area to another.[9] Nearly all manufacturing industries are represented in one or more wage areas, while the metal-working industries, important in war production, have been covered very generally. The inclusion of nonmanufacturing industries has been somewhat more uniform. Retail grocery stores, limited-price variety stores, auto-storage and repair establishments, and financial institutions have been covered in nearly

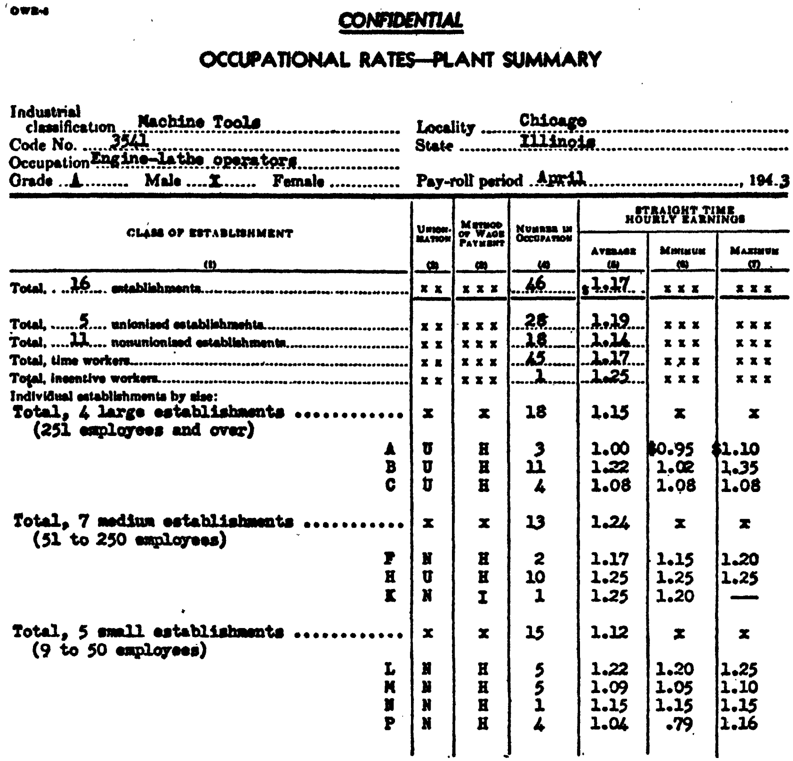
---

[8] During the early months of operation, all facilities of the project were needed to meet the requirements of the War Labor Board. In the late summer of 1943 a series of public releases was undertaken, to provide occupational wage rates by individual industry and locality. Arrangements were also made for the informal release of tabulated data to cooperating companies, unions, and other interested parties.

[9] A few industries, in which wages are determined on a national scale or which are subject to other special considerations, have been specifically excluded from the scope of this project for the time being. Prominent examples are coal mining, the railroads, iron and steel, shipbuilding, aircraft, and building construction. Several of these industries have been subject to recent intensive industry-wide surveys by the Bureau of Labor Statistics.

TABLE I

SAMPLE OF OCCUPATIONAL WAGE RATES FOR ONE OCCUPATION



OWR-6

CONFIDENTIAL

OCCUPATIONAL RATES—PLANT SUMMARY

Industrial classification ..Machine Tools..........
Code No. ....3541..........
Occupation Engine-lathe operators..........
Grade ..A........ Male ....X...... Female ....
Locality ....Chicago..........
State ....Illinois..........
Pay-roll period ..April.........., 194.3

Class of establishment (1)	Unionization (2)	Method of wage payment (3)	Number in occupation (4)	Straight time hourly earnings: Average (5)	Straight time hourly earnings: Minimum (6)	Straight time hourly earnings: Maximum (7)
Total, ..16.... establishments	x x	x x x	46	$1.17	x x x	x x x
Total, ...5.... unionized establishments	x x	x x x	28	1.19	x x x	x x x
Total, ...11.... nonunionized establishments	x x	x x x	18	1.14	x x x	x x x
Total, time workers	x x	x x x	45	1.17	x x x	x x x
Total, incentive workers	x x	x x x	1	1.25	x x x	x x x
Individual establishments by size:						
Total, 4 large establishments .... (251 employees and over)	x	x	18	1.15	x	x
A	U	H	3	1.00	$0.95	$1.10
B	U	H	11	1.22	1.02	1.35
C	U	H	4	1.08	1.08	1.08
Total, 7 medium establishments .... (51 to 250 employees)	x	x	13	1.24	x	x
F	N	H	2	1.17	1.15	1.20
H	U	H	10	1.25	1.25	1.25
K	N	I	1	1.25	1.20	—
Total, 5 small establishments .... (9 to 50 employees)	x	x	15	1.12	x	x
L	N	H	5	1.22	1.20	1.25
M	N	H	5	1.09	1.05	1.10
N	N	H	1	1.15	1.15	1.15
P	N	H	4	1.04	.79	1.16

Note: The sample tabulation presented above is based on a study of 16 representative establishments manufacturing machine tools in the Chicago area. Six of these establishments (D, E, G, I, J, and O) employed no male, grade A, engine lathe operators whose duties were consistent with the Bureau's definition. All of the workers received straight hourly rates except the one worker in establishment K, who was paid on an incentive basis ("I"). The minimum rate shown in column 6 for this incentive worker is his *guaranteed* rate. Other figures appearing in columns 6 and 7 represent the minimum and maximum rates actually paid to individual experienced workers in the respective establishments.

The key jobs used to represent the machinery industries are listed in full in Table II. Tabulations similar to the above are prepared routinely for each key job by grade and sex. In addition, an industry summary (no example shown) is prepared in order to reveal the relationship among average wage rates in the various key jobs; this industry summary presents average rates for all key jobs for significant groups of plants, but does not give detail for individual establishments.

all wage areas, while a more or less standard group of seven additional nonmanufacturing industries has been covered in places of 100,000 population or more. Except in a few industries, establishments with fewer than nine wage earners have not been studied. Otherwise the study has commonly included every establishment in each industry-area unit selected. Sampling of establishments has been practicable on some occasions.

Key Jobs.—Limitation of wage analysis to a few key jobs in each industry has contributed greatly to the simplicity and orderliness of the project and has hastened the processes of collection and tabulation. "Job patterns," consisting of groups of key jobs within individual industries, have been prepared for 120 separate industries. In preparing these patterns the Bureau has relied heavily on the assistance of labor unions, employers, and other groups, as well as on its own experienced field staff.

TABLE II

SAMPLE JOB PATTERN

(Subject to revision)

KEY JOBS IN THE MACHINERY INDUSTRIES (OTHER THAN ELECTRICAL)

Standard Industrial Code No. 35

1. *Maintenance*
    - Carpenters, Grade A and B
    - Electricians, Grade A and B
    - Machinists, Grade A and B
    - Millwrights, Grade A and B

2. *Supervision*
    - Working foremen, processing departments

3. *Processing*
    - Assemblers, bench, Grade A, B, and C
    - Drill-press operators, single spindle, Grade A, B, and C
    - Drill-press operators, multiple spindle, Grade A, B, and C
    - Engine-lathe operators, Grade A, B, and C
    - Milling-machine operators, Grade A, B, and C

3. *Processing (continued)*
    - Screw-machine operators, automatic, Grade A, B, and C
    - Screw-machine operators, hand, Grade A, B, and C
    - Tool and die makers, Grade A and B

4. *Inspection and Testing*
    - Inspectors, Grade A, B, and C

5. *Recording and Control*
    - Stock clerks

6. *Materials Movement*
    - Truck drivers I
    - Truckers, hand
    - Truckers, power

7. *Custodial*
    - Guards
    - Janitors
    - Watchmen

Note: The pattern for the machinery industries is often used in conjunction with other patterns—for example, the patterns for shops manufacturing stamped and pressed metal products, foundries, forge shops, and offices.

The job patterns for nonmanufacturing industries typically include only 10 to 15 key jobs, the selection of which presents few difficulties. Job patterns in manufacturing industries frequently include more than 20 key jobs, which may be selected from a total of 200 or more distinct occupations. Most job patterns cover 40 per cent or more of all wage earners.

In preparing job patterns for manufacturing industry, the labor force is regarded as consisting of seven functional categories: (1) maintenance and repair of plant and equipment; (2) plant supervision; (3) processing; (4) inspection and testing of product; (5) recording and control of raw materials and the flow of work; (6) materials movement; and (7) custodial service. With rare exceptions, each of these categories is represented by one or more jobs in each industry. Thus all major functional groups are represented. The pattern for the machinery industries (other than electrical), presented below for illustrative purposes, includes 21 key jobs, most of which are subdivided by grade.

Except for the processing occupations, these categories are composed of jobs that recur from industry to industry. Thus maintenance jobs are found in nearly all industries, as are working foremen, timekeepers, hand truckers, janitors, and so forth. Such jobs provide a fairly dependable basis for inter-industry comparisons. Inspectors are also found in most industries but are specialized by product and not comparable from one industry to another.

The processing jobs vary widely from industry to industry, reflecting the variety of processes entailed in the manufacture of different products. Because most employees in manufacturing are engaged in processing and because a different set of processing jobs must be selected for almost every industry, these jobs constitute a large proportion of the individual occupations defined and studied in manufacturing industries. The following tabulation[10] indicates the number of different occupations (in manufacturing industries) for which definitions had been prepared by midsummer, 1943:

Maintenance	19
Supervision	2
Processing	634
Inspection	40
Recording	6
Materials movement	23
Custodial	4
Total	728

Within the seven functional categories, individual key jobs are selected by the application of five criteria: These are: (1) definiteness and clarity of the job; (2) numerical importance; (3) critical importance in the war effort; (4) traditional use in wage negotiations; and (5) representativeness of the entire range of wage levels in the industry.

Problems.—A number of troublesome problems have been encountered during the brief life of the Occupational Wage Rate Project. The difficulty of training several hundred field representatives in a period of a few weeks has been mentioned. Obviously the accuracy and consistency of the statistics produced are dependent upon the quality of the field work. Another early problem was the difficulty of obtaining complete universes of establishments classified by city, number of employees, and current product.

---

[10] This tabulation is based on occupations and makes no allowance for subdivisions by skill grade.

The determination of prevailing wage rates in the small community continues to present difficulties. Cities of 25,000 population or less usually contain few manufacturing establishments and these are usually scattered among a number of industries. Tabulations for single industries consequently may yield somewhat unstable results, and may threaten disclosure of individual operations. The alternatives appear to be combination of the data—at least for certain occupations—for different industries within the locality or establishment of fairly broad wage areas including a number of small communities. Although neither of these alternatives is appealing, the latter appears generally to yield the more useful results. Grouping of industries, however, is sometimes unavoidable.

It is, of course, desirable to obtain wage data from all industries and communities as of the same period of time, but this goal is difficult of accomplishment. Assuming that a complete survey of all selected establishments is practicable only once annually,[11] it is apparent that use of the same pay-roll period must involve serious waste in the use of field staff after a peak period, or, if the collection of wage material is spread throughout the year, that much of the wage material will be many months old by the time it is collected. Under the circumstances, there is much to be said for the staggering of pay-roll dates by groups of related industries.

CONCLUSION

In summary, the outstanding features of the Occupational Wage Rate Project appear to be the following: (1) dependence on voluntary cooperation by employers; (2) use of current hourly wage rates as the basic unit of measurement; (3) employment of uniform, written definitions of occupations; (4) particularization of occupational rates by industry and locality; (5) collection of all data from basic records and by means of field visits by trained representatives; (6) limitation of observations primarily to major cities, and to characteristic industries and key jobs; and (7) presentation of findings in detailed but highly standardized form.

These features, although seemingly essential in view of the nature and scale of the stabilization program, have involved certain sacrifices. The selectivity of coverage, for example, has necessitated much interpolation by the administrative agency. The standardized form of presentation dictated by the scale of the project has tended to limit the opportunities for special handling of problem industries or areas. The segregation of wage data by individual industry has made necessary a certain amount of grouping of industry summaries for administrative use. Nevertheless, the project has constituted the chief source of information on prevailing wages and has undoubtedly played an important part in the stabilization program.

---

[11] This does not suggest, of course, that more frequent studies in a small sample of plants may not be necessary for indicating trends of wage rates and for other special purposes.

The adaptation of the project to the emergency needs of the War Labor Board has, of necessity, limited its usefulness for certain other purposes. Representation of the various industries and regions, for example, is by no means uniform, and the heavy demands of the regional Boards have temporarily impeded the broad distribution of the material and limited its general analysis.

In view of the origin of the project as an instrument of wartime control, its post-war future is uncertain. There is little doubt, however, that current information on occupational wage rates will be greatly needed in effecting the transfer of millions of workers from wartime to peacetime jobs. Moreover, the project provides the basis for a long-needed index of wage rates, to supplement present measures of trends in hourly and weekly earnings. Undoubtedly the greatest need for occupational wage rates, however, is for use in wage negotiations; detailed and impartial wage information can contribute substantially to the maintenance of industrial peace.

THE DISTRIBUTION OF DEPOSITS AND CURRENCY IN THE UNITED STATES, 1929–1939

BY SOLOMON SHAPIRO
U. S. Bureau of Labor Statistics

IN SPITE OF THE PLETHORA of statistical data gathered by government agencies and private institutions, there still remain certain serious gaps in the quantitative economic picture. One of these gaps is the lack of detailed information regarding the holdings of money, time deposits, and other liquid savings in the hands of various economic groups. Current information of this nature should help considerably in the adoption of adequate fiscal policies and would contribute to the advancement of monetary theory.

The need for such data is well recognized and some steps are now being taken to obtain them, at least partially, at regular intervals.[1] But little of such information has been available for the period preceding the present war. It is the purpose of this note to present certain data on the distribution of deposits and currency during that period. The accompanying tables furnish an approximation of the amounts of deposits and currency for the years 1929–1939 in the categories indicated.

These estimates are an extension of a study initiated by the Federal Reserve Board.[2] The principal procedures followed in developing the estimates are described in the *Federal Reserve Bulletin* for October 1939. However, certain modifications of the original procedures have been deemed desirable. The most important of these are:

(1) Adjustments for demand deposits were made only for items in transit between banks, whereas the Federal Reserve Board study made an additional adjustment for items in transit between depositors. Without the second adjustment the latter "float" becomes part of the unclassified holdings.

(2) The holdings of corporations in demand deposits, time deposits, and currency were estimated on the basis of the proportion of these components of total cash in the various industries as found by the Federal Reserve Board in its study of large deposit accounts for 1933

---

[1] The Federal Reserve Board, in cooperation with the reserve banks, is planning to survey a considerable number of banks every six months to classify deposits by type of holder. See *Federal Reserve Bulletin*, August 1943. The Securities and Exchange Commission's periodic releases giving estimates of the components of individual savings are also of value in meeting this need.

[2] The procedures used in the Federal Reserve Board study were developed by Martin Krost, who compiled estimates for the end of 1933 and 1935. Lauchlin Currie presented a paper based on these estimates before the American Statistical Association in December 1937 (published in this JOURNAL for June 1938). Haskell Wald, using the techniques developed in the earlier study, made estimates for the end of 1936 and 1937, published in the *Federal Reserve Bulletin* for October 1939 and May 1940, respectively.

and 1935. The proportions prevailing in 1933 were applied to cash holdings in 1929, 1930, 1931, 1932, and 1933. The 1935 proportions were used for the years 1935, 1936, 1937, 1938, and 1939. For the year 1934, the averages of 1933 and 1935 were used.

(3) Holdings of security brokers, assumed to be demand deposits only, were estimated by the Federal Reserve Board for 1933 and 1935 through 1939. The approximate relationship during these years between brokers' borrowings by New York Stock Exchange members and their money holdings was the basis for deriving the estimates of the earlier years.

(4) Deposits of banks' trust departments have been included with the unclassified group. This was done because estimates of these deposits, based on the proportion of such deposits in national banks to similar deposits in insured banks in May 1936, become quite unreliable for the period prior to 1933 when the number of national banks exercising trust powers was considerably larger.

(5) Deposits of other public bodies were estimated through 1935 on the basis of the percentage change between June and December of such deposits in Federal Reserve member banks, applied to similar deposits in all banks on June 30. From 1936 such deposits are reported by the Comptroller of the Currency for year-end dates.

The reliability of the data varies from those figures regularly reported, such as United States Treasury holdings, to estimates based on rather meager evidence, such as holdings of security brokers. Reliability likewise varies for the different years depending on the adequacy of the basic data from which the estimates were derived. On the whole, the data for the distribution of deposits and currency, except for the holdings of non-corporate enterprises and security brokers, present a satisfactory degree of reliability. The distributions of demand deposits and currency, and demand deposits only, become less accurate since they require the deduction of estimates of time deposits, and time deposits and currency, respectively, for the various holders. Nevertheless, even the figures whose accuracy may be questioned should be decidedly useful for general indications of the various magnitudes.

A detailed analysis of the underlying reasons for the changes in the money supply (demand deposits and currency) and the shifts in its distribution would shed a great deal of light on the economic tendencies in operation during the period. Such an analysis of the data is beyond the scope of this note, but a few significant changes indicated by the figures for the money supply and the important economic groups of money holders, Table I, may be noted.

Changes in the magnitude of the money supply during the decade of

TABLE I

THE DISTRIBUTION OF DEMAND DEPOSITS AND CURRENCY BY TYPE OF HOLDER

(amounts in millions of dollars—end of year figures)

	1929		1930		1931		1932		1933		1934		1935		1936		1937		1938		1939	
	Amount	Per cent	Amount	Per cent	Amount	Per cent	Amount	Per cent	Amount	Per cent	Amount	Per cent	Amount	Per cent	Amount	Per cent	Amount	Per cent	Amount	Per cent	Amount	Per cent
*Total**	*27,290*	*100.0*	*25,540*	*100.0*	*23,000*	*100.0*	*21,270*	*100.0*	*20,980*	*100.0*	*25,180*	*100.0*	*29,060*	*100.0*	*32,860*	*100.0*	*31,840*	*100.0*	*34,360*	*100.0*	*38,880*	*100.0*
Business—total	*8,780*	*32.2*	*8,570*	*33.5*	*7,050*	*30.7*	*6,900*	*32.4*	*6,470*	*30.9*	*7,240*	*28.8*	*7,830*	*27.0*	*8,800*	*27.0*	*8,080*	*25.8*	*9,070*	*26.4*	*9,950*	*25.6*
Corporations engaged in:																						
Manufacturing, mining, and construction	3,800	14.0	3,820	14.9	3,260	14.2	3,120	14.7	2,900	13.8	3,030	12.0	3,530	12.1	3,800	11.6	3,550	11.3	4,200	12.3	4,800	12.3
Transportation and public utilities	1,340	4.9	1,390	5.4	1,070	4.7	1,090	5.1	1,080	5.2	1,120	4.5	1,140	3.9	1,450	4.4	1,080	3.5	1,320	3.8	1,450	3.7
Trade, service, agriculture, and miscellaneous	1,730	6.3	1,550	6.1	1,270	5.5	1,270	6.0	1,200	5.7	1,530	6.1	1,560	5.4	1,790	5.5	1,690	5.4	1,730	5.0	1,780	4.6
Non-corporate enterprises†	1,910	7.0	1,810	7.1	1,450	6.3	1,420	6.6	1,290	6.2	1,570	6.2	1,610	5.6	1,820	5.5	1,750	5.6	1,820	5.3	1,920	5.0
Finance—total	*5,570*	*20.4*	*4,480*	*17.3*	*2,930*	*12.7*	*2,330*	*11.0*	*2,020*	*9.6*	*2,820*	*11.2*	*3,960*	*13.6*	*4,230*	*12.8*	*4,160*	*13.3*	*4,790*	*13.9*	*6,360*	*16.4*
Corporations engaged in:																						
Insurance	400	1.5	370	1.4	350	1.5	470	2.2	590	2.8	760	3.0	940	3.2	990	3.0	930	3.0	1,060	3.1	1,310	3.4
Other finance	2,050	7.5	1,900	7.4	1,240	5.4	1,050	5.0	860	4.1	1,350	5.4	1,770	6.1	1,630	4.9	1,550	4.9	1,740	5.1	1,960	5.0
Non-corporate:																						
Security brokers	1,820	6.7	1,010	4.0	370	1.6	100	.5	200	1.0	260	1.0	200	.7	300	.9	280	.9	240	.7	260	.7
Foreign banks	550	2.0	550	2.2	480	2.1	260	1.2	130	.6	170	.6	470	1.6	530	1.6	630	2.0	700	2.0	1,150	3.0
Other foreigners‡	750	2.7	590	2.3	490	2.1	450	2.1	240	1.1	290	1.2	570	2.0	780	2.4	780	2.5	1,050	3.0	1,680	4.3
Public bodies—total	*1,890*	*6.9*	*1,970*	*7.7*	*2,060*	*8.9*	*1,810*	*8.5*	*2,550*	*12.2*	*4,080*	*16.2*	*4,170*	*14.3*	*4,170*	*12.7*	*3,620*	*11.6*	*4,820*	*14.0*	*4,430*	*11.4*
United States Treasury	190	.7	330	1.3	520	2.3	520	2.4	1,020	4.9	1,840	7.3	1,450	5.0	1,240	3.8	970	3.1	1,820	5.3	1,480	3.8
Other public bodies	1,700	6.2	1,640	6.4	1,530	6.6	1,290	6.1	1,530	7.3	2,240	8.9	2,710	9.3	2,930	8.9	2,650	8.5	3,000	8.7	2,930	7.6
Unclassified—balances of individual, not-profit organizations, banks' trust departments, etc.	*11,060*	*40.5*	*10,580*	*41.5*	*10,970*	*47.7*	*10,230*	*48.1*	*9,940*	*47.3*	*11,030*	*43.8*	*13,110*	*45.1*	*15,600*	*47.5*	*15,480*	*49.3*	*15,690*	*45.7*	*18,100*	*46.6*

* Demand deposits are adjusted for items in process of collection by banks and exclude domestic interbank deposits.

† Excludes professional service, agriculture, finance, transportation, and public utilities.

‡ Includes foreign bank deposits with non-member banks.

Note: Amounts are rounded and will not necessarily add up to totals.

TABLE II

THE DISTRIBUTION OF TOTAL DEPOSITS AND CURRENCY BY TYPE OF HOLDER

(millions of dollars—end of year figures)

	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939
*Total**	*55,470*	*54,150*	*48,990*	*45,710*	*42,770*	*48,370*	*53,290*	*58,260*	*57,600*	*60,740*	*65,880*
Business—total	*10,230*	*10,000*	*8,230*	*8,050*	*7,560*	*8,080*	*8,420*	*9,540*	*8,670*	*9,740*	*10,700*
Corporations engaged in:											
Manufacturing, mining, and construction	4,560	4,580	3,910	3,750	3,470	3,450	3,810	4,110	3,840	4,540	5,190
Transportation and public utilities	1,680	1,740	1,350	1,360	1,360	1,320	1,260	1,610	1,200	1,460	1,600
Trade, service, agriculture, and miscellaneous	1,850	1,660	1,360	1,370	1,290	1,620	1,650	1,689	1,790	1,820	1,880
Non-corporate enterprises†	2,140	2,010	1,610	1,570	1,430	1,690	1,700	1,930	1,840	1,920	2,030
Finance—total	*6,200*	*5,160*	*3,170*	*2,530*	*2,240*	*3,100*	*4,250*	*4,530*	*4,470*	*5,130*	*6,760*
Corporations engaged in:											
Insurance	480	450	430	570	720	920	1,150	1,210	1,130	1,290	1,590
Other finance	2,230	2,060	1,350	1,140	930	1,440	1,850	1,700	1,620	1,820	2,050
Non-corporate:											
Security brokers	1,820	1,010	370	100	200	260	200	300	280	240	260
Foreign banks	700	790	500	260	140	170	480	540	640	710	1,160
Other foreigners‡	960	850	520	450	250	310	580	790	800	1,070	1,710
Public bodies—total	*2,450*	*2,490*	*2,450*	*2,170*	*2,860*	*4,390*	*4,580*	*4,560*	*4,270*	*5,410*	*4,950*
United States Treasury	190	330	520	520	1,020	1,840	1,450	1,240	970	1,820	1,480
Other public bodies	2,270	2,160	1,930	1,650	1,840	2,550	3,120	3,320	3,290	3,580	3,470
Unclassified—balances of individuals, non-profit organisations, banks' trust departments, etc.	*36,600*	*36,510*	*35,140*	*32,970*	*30,110*	*32,800*	*36,040*	*39,630*	*40,190*	*40,460*	*43,470*

* Demand deposits are adjusted for items in process of collection by banks and exclude domestic interbank deposits.

† Excludes professional service, agriculture, finance, transportation, and public utilities.

‡ Includes foreign bank deposits with non-member banks.

Note: Figures are rounded and will not necessarily add up to totals.

TABLE III

THE DISTRIBUTION OF DEMAND DEPOSITS BY TYPE OF HOLDER

(millions of dollars—end of year figures)

	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939
*Total**	*23,720*	*21,950*	*18,500*	*16,560*	*16,180*	*20,510*	*24,120*	*27,310*	*25,650*	*28,530*	*32,420*
Business—total	*8,340*	*8,120*	*6,690*	*6,540*	*6,130*	*6,870*	*7,450*	*8,430*	*7,680*	*8,630*	*9,480*
Corporations engaged in:											
Manufacturing, mining, and construction	3,680	3,700	3,160	3,020	2,800	2,940	3,430	3,690	3,450	4,080	4,660
Transportation and public utilities	1,220	1,260	980	990	980	1,030	1,050	1,340	1,000	1,220	1,340
Trade, service, agriculture, and miscellaneous	1,620	1,450	1,180	1,190	1,120	1,430	1,460	1,680	1,590	1,610	1,660
Non-corporate enterprises:†	1,830	1,710	1,370	1,340	1,220	1,480	1,520	1,720	1,650	1,720	1,810
Finance—total	*5,470*	*4,330*	*2,860*	*2,270*	*1,960*	*2,740*	*3,850*	*4,120*	*4,060*	*4,670*	*6,220*
Corporations engaged in:											
Insurance	380	350	340	440	560	720	900	950	890	1,010	1,250
Other finance	1,970	1.830	1,190	1,010	820	1,300	1,700	1,560	1,490	1,670	1,880
Non-corporate:											
Security brokers	1,820	1,010	370	100	200	260	200	300	280	240	260
Foreign banks	550	550	480	260	130	170	470	530	630	700	1,150
Other foreigners‡	750	590	490	450	240	290	570	780	780	1,050	1,680
Public bodies—total	*1,890*	*1,970*	*2,060*	*1,810*	*2,550*	*4,080*	*4,170*	*4,170*	*3,620*	*4,820*	*4,420*
United States Treasury	190	330	520	520	1,020	1,840	1,450	1,240	970	1,820	1,480
Other public bodies	1,700	1,640	1,530	1,290	1,530	2,240	2,710	2,930	2,650	3,000	2,930
Unclassified—balances of individuals, non-profit organizations, banks' trust departments, etc.	*8,020*	*7,530*	*6.900*	*5,940*	*5,540*	*6,820*	*8,660*	*10,590*	*10.280*	*10,410*	*12,310*

* Demand deposits are adjusted for items in process of collection by banks and exclude domestic interbank deposits.

† Excludes professional service, agriculture, finance, transportation, and public utilities.

‡ Includes foreign bank deposits with non-member banks.

Note: Figures are rounded and will not necessarily add up to totals.

TABLE IV

THE DISTRIBUTION OF TIME DEPOSITS BY TYPE OF HOLDER

(millions of dollars—end of year figures)

	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939
Total	*28,120*	*28,610*	*25,990*	*24,440*	*21,790*	*23,190*	*24,230*	*25,400*	*26,260*	*26,380*	*27,060*
Business—total	*1,450*	*1,430*	*1,180*	*1,150*	*1,080*	*840*	*590*	*680*	*600*	*680*	*750*
Corporations engaged in:											
Manufacturing, mining, and construction	760	760	650	620	580	420	290	310	290	340	390
Transportation, and public utilities	340	350	270	280	270	200	130	160	120	140	150
Trade, service, agriculture, and miscellaneous	130	110	100	90	90	100	90	100	90	90	100
Non-corporate enterprises:*	230	200	160	160	140	130	90	110	100	110	110
Finance—total	*630*	*740*	*240*	*190*	*220*	*280*	*300*	*310*	*300*	*340*	*410*
Corporations engaged in:											
Insurance	90	80	80	100	130	170	210	220	200	230	290
Other finance	180	160	110	90	70	90	80	70	70	80	90
Non-corporate:											
Security brokers	—	—	—	—	—	—	—	—	—	—	—
Foreign banks	150	240	30	—	10	10	10	10	10	10	10
Other foreigners†	210	260	30	—	10	10	10	10	20	20	20
Public bodies—total	*570*	*520*	*390*	*360*	*310*	*310*	*410*	*390*	*640*	*580*	*530*
United States Treasury	—	—	—	—	—	—	—	—	—	—	—
Other public bodies	570	520	390	360	310	310	410	390	640	580	530
Unclassified—balances of individuals, non-profit organizations, banks' trust departments, etc.	*25,540*	*25,920*	*24,180*	*22,740*	*20,170*	*21,770*	*22,930*	*24,030*	*24,710*	*24,780*	*25,370*

* Excludes professional service, agriculture, finance, transportation, and public utilities.

† Includes foreign bank deposits with non-member banks.

Note: Figures are rounded and will not necessarily add up to totals.

the thirties reflect the trends of business activity, sharply downward from 1929 until 1933 and upward thereafter. Generally speaking, all the groups and subgroups of holders follow the movements indicated by the aggregate, although the rise in the holdings of insurance companies appears to lead the upward movement.

At the end of 1939 the business group held over a billion dollars more than at the beginning of the decade. Nevertheless, the proportion of the total money supply held by business declined throughout the period, from roughly one-third of the aggregate in 1929 to about one-fourth in 1939. The relative importance of the financial group likewise declined, from approximately one-fifth to one-sixth of the total. This relative decrease occurred in spite of substantial additions to the holdings of insurance companies and foreigners toward the end of the thirties. Public bodies began to add to their money holdings after 1932, largely due to a rather steady growth of state and local government funds. United States Treasury holdings fluctuated considerably but at a higher level after the early years. Some of the importance lost by the business and finance groups was gained by public bodies in the shifts occurring over the period. The unclassified group of holders increased their money both in absolute and in relative terms during the decade. This group being residual in its derivation is composed not only of consumers and non-profit organizations, but also of elements which are really financial in nature, such as banks' trust departments and very large accounts of individuals. The explanation of the changes which occurred in this last category will therefore be even more complex than the causes of shifts affecting the other groups.

VARIATIONS IN EXPENDITURES WHERE FAMILIES OF WAGE EARNERS AND CLERICAL WORKERS ARE CLASSIFIED BY ECONOMIC LEVEL

BY ROY C. CAVE

SINCE THE APPEARANCE of the book by Allen and Bowley on *Family Expenditures*,[1] most such studies have dealt with expenditures of families classified according to net family income or total family expenditure, not according to total expenditure per expenditure unit.[2] In the present study, analyses are made of variations of expenditures at 12 economic levels (1) on selected main categories of items by 446 San Francisco-Oakland families sorted by economic level; and (2) per person on selected individual clothing items by 14,469 families, sorted by economic level. The basic data used are taken from the nationwide survey, "Money Disbursements of Wage Earners and Clerical Workers, 1934–36," made by the Bureau of Labor Statistics of the United States Department of Labor.

A. D. H. Kaplan has already shown that, in Chicago, for most general categories or groups of commodities, use of linear regressions is feasible where families are classified by income and occupation, up to the $5,000 level.[3] One purpose of the present study is to indicate the feasibility of the use of linear regressions where families with net incomes below $4,000, and belonging to related occupational groups, are classified by economic level. The first of the two analyses made illustrates the use of the least squares method where family expenditures on groups of items are plotted against total family expenditures. The second illustrates the same method where expenditures by individuals on specific items are plotted against expenditures on the sub-group to which those specific items belong.[4]

---

[1] Allen, R. D. G., and Bowley, A. L., *Family Expenditures, A Study of Its Variation*, 1935.

[2] When families are classified by income level, no account is taken of the effect of differences in family size and composition on the plane of living. To meet this difficulty, families may be classified by economic level, as measured by the amount spent per total expenditure unit. The Bureau of Labor Statistics calculates the total expenditure unit for any given family by combining the amounts spent by the family per food expenditure unit and per clothing expenditure unit, and adding to this sum the amount the family spends per adult male equivalent for "other items." The food expenditure unit is the average expenditure for food of full-time employed men, aged 20 and over. The clothing expenditure unit is the average annual clothing expenditure of white men wage earners and clerical workers aged 21 to 35 inclusive, in all cities covered by the 1934–36 expenditure study of the Bureau. *Cf.* U. S. Bureau of Labor Statistics, Bulletin No. 639, 1939, pp. 344–351.

[3] Kaplan, A. D. H., "Expenditure Patterns of Urban Families," this JOURNAL, Vol. 33, No. 201, March, 1938, pp. 88–100.

[4] By means of the method of least squares it is possible to express broad consumption tendencies, such as Engel's law of food consumption. The constant *b* indicates the slope of the regression line, and at the same time the dollar (or fraction thereof) increases in expenditure on a given item for every dollar increase in spendable funds.

Another purpose is to provide additional bases for estimating family and individual expenditures at different economic levels, for classifying groups of items and individual items according to urgency of wants for them, and for estimating elasticity of demand.[5] Various coefficients and other measures have been computed to indicate accuracy of estimates, degree of correlation, and expenditure elasticity.

At low and modest income levels, total expenditure provides a more suitable basis for many purposes than net family income for classifying families in terms of *effective* incomes, since total expenditure tends to exceed net income. And when families are classified in terms of total expenditure per expenditure unit, the influence of increasing effective income on group or individual expenditures can be measured still more precisely. This is especially true if the study is confined to a given occupational group, or closely related groups. Expenditure data analyzed here are taken from two groups (families of wage earners and clerical workers) having about the same income range and similar habits of expenditure. Where wage earner groups are analyzed separately from business and professional groups, many types of expenditure tend to become more linear, since at higher income levels expenditures increase more slowly, and thus at upper levels expenditure curves covering a wide range of income will turn downward.

Expenditures of the 446 San Francisco-Oakland wage earner and clerical families, when classified by income rather than economic level, were estimated by Williams and Hanson[6] from the regression equation $Y=a+bX^{1/2}+cX$, where X is annual net income. This equation was selected after experimentation had established that most categories of expenditure were curvilinear.[7] It was found to give about as good a fit for those categories which were linear. Values of a, b, and c are not given with estimated expenditures.

When classified by economic level, as might be expected, these same families have average expenditures which tend to be more linear. Food, housing, and fuel, light, and refrigeration, all of which were found by Williams and Hanson to be curvilinear when the data are sorted by income, are definitely linear, while clothing outlays are only slightly curvilinear. Regressions for only these four main categories are shown in the present analysis. However, additional estimates have been made for other household operation; furnishings and equipment; automobiles, motorcycles and other transportation; and all other items. Only the category, automobiles, motorcycles and other transportation tends to be curvilinear.[8]

---

[5] If the group of families analysed is a representative sample of the universe to which it belongs, and if further no material change has occurred in retail prices during the period of the study, it may be assumed that an increase of funds available for expenditure is roughly equivalent to a decrease of price with spendable funds remaining unchanged. Hence, measures of expenditure elasticity are also approximations to demand elasticity.

[6] Williams, F. M., and Hanson, A. C., "Money Disbursements of Wage Earners and Clerical Workers in Five Cities of the Pacific Region, 1934–36." U. S. Bureau of Labor Statistics Bulletin 639, 1939, pp. 284, 304–305.

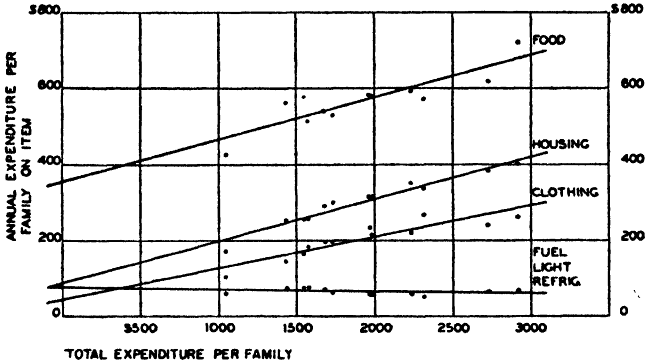
[7] Only personal care, medical care, household operation other than fuel, light, and refrigeration, gifts, and miscellaneous items were truly linear. Weighting by the number of cases, total expenditure, etc., did not yield better results than use of the average actual expenditure for each income group.

In Chart I, annual expenditures per family (Y) for four groups of items are plotted against total expenditure per family (X) for twelve economic levels, and a line of least squares is fitted to each series of data.

CHART I

SAN FRANCISCO-OAKLAND WHITE, NON-RELIEF WORKERS* FAMILIES, 1934–36

Expenditures of 446 Families, Classified by Economic Level† for Food, Housing, Clothing, and Fuel, Light, and Refrigeration



* Derived from U. S. Bureau of Labor Statistics, Bulletin 639, "Money Disbursements of Wage Earners and Clerical Workers in Five Cities in the Pacific Region, 1934–36," 1939, p. 132.

† That is, families are classified according to their total expenditures per total expenditure unit, not according to total income per family without regard to family size and composition.

At the zero point of origin a is positive in the four cases, indicating that all are "necessities." In order of urgency, that complex of wants known as food ranks first, followed in turn by the complexes known as housing, fuel, light and refrigeration, and clothing. While this ranks food as the prime necessity, two qualifications seem required. First, food is not simply a necessity, but a complex of wants, some of which are comforts and luxuries. The same is true of clothing, housing, and other groups of items. The second qualification is that extending the fitted line beyond the plotted data is a questionable procedure.

---

[8] The following equations were obtained:

other household operation—$Y = 1.00 + .05033X$
furnishings and equipment—$Y = -57.17 + .0692X$
all other items—$Y = -127.1 + .2515X$

Note that b is slightly negative in the case of fuel, light, and refrigeration. This negative tendency is partly due to the fact that nearly one-half of the 446 families lived in apartments, and that some of these latter made no separate payment for heat. A larger proportion of families in high than those at low income levels lived in apartments, and these families were especially concentrated at high income levels. Families living in apartments spend less on the average for fuel, light and refrigeration even when they pay for heat. Families at higher levels, too, tend to be smaller in size, so that fuel, light and refrigeration requirements are lower. Among families making no such separate payment, those at the highest levels apparently did less cooking at home.[9] Only the data for clothing show any definite signs of being curvilinear.

Table I gives essential estimates and measures for the four groups of items. The standard error of estimate (S_y) and the coefficient of correlation (r_{yx}) are adjusted for the number of cases (i.e., the number of averages). The coefficient, and adjusted coefficient of correlation are negative in the case of fuel, light, and refrigeration. They are also so low as to indicate very little relationship, for this group of items, between total expenditure and group expenditure. Evidently, too, clothing outlays, while affected by rising economic level, are substantially influenced as well by other considerations. The adjusted coefficient of determination ($\bar{d}_{yx}$) is only .479. Note the high correlation and coefficient of determination in the case of housing expenditures. The standard error[10] of b_{yx} is relatively higher in the cases of clothing, and fuel, light, and refrigeration. In the case of the latter, it is almost as large as b, but b is exceedingly small. The expenditure elasticity (η) is considerably less than unity in all cases.[11] It is negative and extremely low for fuel, light, and refrigeration, very low for food, and highest for clothing. Elasticities for food, clothing, and fuel, light, and refrigeration decline, and housing elasticity increases, when families are grouped by economic level, judging by comparisons made with Kaplan's Chicago study in which families are grouped by income. Stated differently, when families are arranged according to effective income, i.e., funds available for ex-

---

[9] Williams and Hanson, *op. cit.*, pp. 211–212. The outlay for gas tended to decline.

[10] Cf. Ezekiel, Mordecai, *Methods of Correlation Analysis*, 2nd ed., 1941, pp. 312–14. The formula for σ_b is $\dfrac{\bar{S}_y}{\sigma_x \sqrt{n}}$.

[11] Computed by the formula, $\eta = k/w$, where η = expenditure elasticity, k = the constant b, and $\bar{w}$ = the percentage of total expenditures devoted by all families to the group of items; or $\eta = b/p$ since $b = \Delta e/\Delta i$ and $e/i = p$, where Δe is increase of expenditure, Δi is increase of total expenditure, and p is percentage of total expenditure.

TABLE I

LINEAR REGRESSION CONSTANTS, ELASTICITY OF EXPENDITURE, AND MEASURES OF ACCURACY OF ESTIMATE AND DEGREE OF CORRELATION

Expenditures of 446 San Francisco-Oakland White Non-Relief Families of Wage Earners and Clerical Workers, Classified by Economic Level, 1934–36

Economic level	Total expenditure per family	Food		Housing		Clothing		Fuel, light, refrigeration	
	X	Y	Y_c	Y	Y_c	Y	Y_c	Y	Y_c
Under $ 300	$1,039	$ 425	$470	$ 171	$ 203	$ 104	$ 130	$ 61	$ 70.22
$ 300– 400	1,429	562	513	251	247	146	162	74	68.66
400– 500	1,541	579	526	255	260	163	171	74	68.22
500– 600	1,570	515	529	256	263	183	174	78	68.10
600– 700	1,686	542	542	290	276	198	183	69	67.64
700– 800	1,723	532	546	304	280	195	186	67	67.49
800– 900	1,977	582	574.6	314	309	239	206.9	64	66.47
900–1,000	1,984	579	575.4	316	310	210	207.4	64	66.44
1,000–1,100	2,232	596	603	353	338	222	228	63	65.45
1,100–1,200	2,311	573	612	339	346	267	234	54	65.14
1,200–1,300	2,722	619	658	388	393	241	268	64	63.49
1,300 and over	2,914	725	680	402	414	266	284	68	62.72
Total	$23,128	$6,829	$6,829	$3,639	$3,639	$2,434	$2,434.3	$800	$801.04
Mean	$ 1,927.3	$ 569.08		$ 303.25		$ 202.8		$ 66.67	

Measures and adjustments				
a_{yx}	$+353.22	$+86.62	$+44.80	$+74.38
b_{yx}	$+ .112	$+ .1124	$+ .082	$− .004
σ_b	$ .0198	$ .0086	$ .0119	$ .0035
S_y	$ 32.6	$ 14.1	$ 19.6	$ 5.9
$\bar{S}_y$	$ 35.7	$ 15.5	$ 21.4	$ 6.4
r_{yx}	+ .875	+ .95	+ .73	− .34
$\bar{r}_{yx}$	+ .861	+ .945	+ .697	− .16
d_{yx}	.766	.903	.53	.116
$\bar{d}_{yx}$	.742	.893	.486	.027
η	.338	.677	.745	− .095
σ_x	$ 521.00			

penditure per expenditure unit, distinctly different elasticities are obtained.

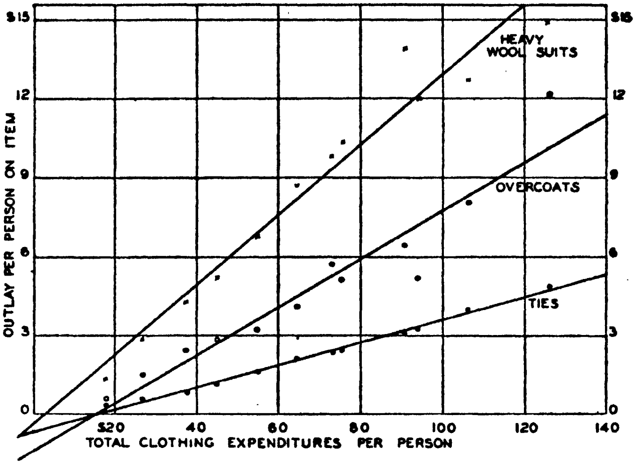
Chart II shows plotted data and fitted lines for six selected individual clothing items. In this case, Y represents expenditure on the individual item and X total clothing outlay per person for men, and boys 18 years of age and over.[12] Since the intercept a is negative for all items except street shoes, it would appear that street shoes alone should be regarded as a necessity. However, this is extension of the fitted line beyond the

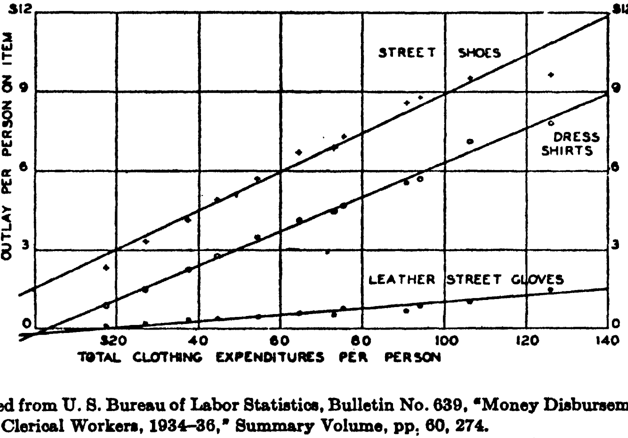
---

[12] X may be regarded as total expenditure for a sub-group, by economic level, since expenditures of men and boys 18 and over constitute a family sub-group.

CHART II

CLOTHING EXPENDITURES* BY ECONOMIC LEVEL† AND PER PERSON, MEN AND BOYS 18 AND OVER,‡ IN 14,469 FAMILIES OF NON-RELIEF WHITE AND NEGRO WAGE EARNERS AND CLERICAL WORKERS, 1934–36, 42 CITIES OF UNITED STATES





---

* Derived from U. S. Bureau of Labor Statistics, Bulletin No. 639, "Money Disbursements of Wage Earners and Clerical Workers, 1934–36," Summary Volume, pp. 60, 274.

† Families are classified according to total expenditure per expenditure unit.

‡ Expenditure per person, with only persons dependent on family funds for 52 weeks included.

data. In point of fact, as total expenditures approach zero, there is a noticeable tendency for practically all plotted expenditures to turn downward. Expenditures for most items become more elastic at very low income or economic levels. At any rate, judging from the height

TABLE II

LINEAR REGRESSION CONSTANTS, ELASTICITY OF EXPENDITURE, AND MEASURES OF ACCURACY OF ESTIMATE AND DEGREE OF CORRELATION

Clothing expenditures by economic level of men, and boys 18 and over, in 14,469 white and Negro families of wage earners and clerical workers in 42 cities of the United States
(Non-Relief Families, Expenditures For One Year During 1934–36)

Economic level	Total clothing expenditure per person	Heavy wool suits		Overcoats		Ties		Street shoes		Dress shirts		Leather street gloves	
	X	Y	Y_c	Y	Y_c	Y	Y_c	Y	Y_c	Y	Y_c	Y	Y_c
Under $ 300	$ 17.42	$ 1.34	$ 1.93	$ .60	$ .22	$ .26	$ .13	$2.34	$2.84	$ .90	$ .97	$.04	$ .00
$ 300– 400	26.89	2.84	3.19	1.51	1.09	.57	.53	3.33	3.53	1.53	1.58	.16	.11
400– 500	37.32	4.31	4.58	2.45	2.04	.93	.97	4.16	4.28	2.28	2.25	.24	.23
500– 600	44.65	5.17	5.56	2.76	2.71	1.26	1.28	4.85	4.82	2.79	2.73	.33	.32
600– 700	54.47	6.80	6.86	3.27	3.61	1.69	1.70	5.70	5.53	3.44	3.36	.44	.44
700– 800	64.06	8.77	8.14	4.19	4.49	2.13	2.11	6.70	6.23	4.14	3.98	.55	.55
800– 900	73.07	9.85	9.34	5.81	5.31	2.43	2.49	6.87	6.88	4.46	4.57	.54	.65
900–1,000	75.87	10.43	9.71	5.17	5.57	2.53	2.61	7.24	7.08	4.73	4.75	.76	.69
1,000–1,100	90.88	13.99	11.71	6.45	6.94	3.09	3.25	8.62	8.17	5.56	5.72	.66	.86
1,100–1,200	93.86	12.04	12.11	5.18	7.21	3.30	3.37	8.82	8.39	5.72	5.91	.88	.90
1,200–1,300	106.14	12.75	13.75	8.11	8.34	3.97	3.90	9.49	9.28	7.15	6.71	1.00	1.04
1,300 and over	125.73	14.95	16.36	12.16	10.13	4.90	4.73	9.62	10.70	7.82	7.97	1.47	1.27
Total	$810.36	$103.24	$103.24	$57.66	$57.66	$27.06	$27.07	$77.74	$77.73	$50.52	$50.50	$7.07	$7.06
Mean	$ 67.53	$ 8.603		$ 4.805		$ 2.255		$ 6.478		$ 4.21		$ .589	
Measures and adjustments													
a_{yx}		$−.392		$−1.369		$−.615		$+1.575		$−.16		$−.204	
b_{yx}		$+.1332		$ +.09143		$+.0425		$ +.0726		$+.0647		$+.01175	
σ_b		$ .009		$ .0089		$ .000885		$ .00423		$ .001658		$ .00092	
S_y		$ .916		$ .896		$ .089		$ .426		$ .167		$ .093	
$\bar{S}_y$		$ 1.003		$ .982		$ .098		$ .467		$ .183		$ .102	
σ_x		$31.827		$ 31.827		$31.827		$ 31.827		$31.827		$31.827	
r_{yx}		+.976		+.955		+.999		+.944		+.996		+.97	
$\bar{r}_{yx}$		+.947		+.95		+.997		+.938		+.991		+.93	
d_{xy}		.95		.912		.998		.89		.992		.94	
$\bar{d}_{xy}$		.896		.902		.994		.88		.982		.865	
η		1.07		1.40		1.43		.69		1.04		1.53	

of a, in order of urgency street shoes rank first, followed by dress shirts, leather gloves, suits, ties and overcoats. It may be that the last five can be made to last longer if income is insufficient, and that gloves and ties, at least, can be eliminated entirely.

Table II gives estimated regressions of Y on X, and summarizes important measures bearing on accuracy of estimate and degree of correlation. For these data b represents the dollar change (or fraction thereof) in expenditure on the item for every dollar increase in average total clothing expenditure by men, and boys 18 years and over. This coefficient, b_{yx}, is largest for heavy wool suits and overcoats, followed next by street shoes and dress shirts. For leather gloves, b is extremely small.

S_y is larger relative to size of estimated expenditures (Y_c) in the case of these expenditures on individual items than in the previous case where expenditures are for groups of items. This illustrates the value of accompanying measures of accuracy. The five "luxury" items all necessarily have an expenditure elasticity greater than unity (since the intercept a is negative in each case), while the ratio for street shoes is only .69. Variation in the expenditure elasticity of even these few individual clothing items demonstrates the inaccuracy of speaking of clothing in general as a "necessity," or a "luxury," depending on the method of income classification.[13]

---

[13] For families grouped by net family income, the general category clothing tends to be a "luxury," since η is greater than 1. For families classified by economic level, it is a "necessity" since η is definitely less than 1.

FITTING POLYNOMIAL TRENDS TO SEASONAL DATA BY THE METHOD OF LEAST SQUARES

By Howard L. Jones

THE PROBLEM discussed in this article is that of fitting a polynomial trend to an equally spaced time series with unknown seasonal differences in such manner that the seasonal differences and the coefficients of the polynomial are computed simultaneously so as to minimize the sum of the squared residuals. This problem is simple in theory; but its solution appears to have been overlooked in the literature on fitting trends.[1] Sasuly[2] suggests that fitting certain arcs by an iterative procedure will tend to eliminate the seasonal element; but there is no reason to suppose that this procedure will approximate the least squares solution. Macaulay[3] argues that the method of least squares is not suitable for computing trend midordinates for a seasonal series, and chooses flexibility and smoothness as principal criteria for judging graduation formulas devised by himself and other writers. His argument seems to rest chiefly on the fact that a moving arc fitted to a monthly series by a method that ignores the seasonal differences is obviously a biased estimate of the trend. In fitting moving arcs by the procedure outlined in Table I, the seasonal differences are taken into account.

TABLE I

WEIGHTS TO BE APPLIED TO TWELVE-MONTH MOVING TOTALS TO FIND THE MIDORDINATE OF A FITTED THIRD-DEGREE PARABOLA

No. of terms, original series	15	17	19	21	23
No. of moving totals	4	6	8	10	12
Multiply sum of first and last totals by	−73	−158	−267	−412	−55
Multiply sum of all intermediate totals by	85	109	145	193	23
Add products and divide by	288	1440	4032	8640	1440
Standard error ratio	.649	.425	.363	.335	.318

Let Y_x be the general expression for one of a series of n terms with subscripts at unit intervals of x ranging from $-l$ to $+l$ where $l=\frac{1}{2}(n-1)$. We are to fit the function

---

[1] In the *Journal of Accountancy* for September, 1937 (Vol. 64, pp. 174–199), Arne Fisher discusses the problem of fitting a fifth-degree polynomial, in the orthogonal form, to 25 terms in a monthly series. His solution, however, is defective in at least two respects: (1) To find the first coefficient of his polynomial, he takes the arithmetic mean of the 25 terms; but this will yield a biased estimate because one of the seasonal differences is contained three times in the series while each of the other seasonal differences is contained but twice. (2) Taking the arithmetic mean to find the first coefficient implies that the observational errors are independent; but the other coefficients are obtained by fitting the 13 differences obtained by subtracting each observation from the observation twelve months later, which implies that the errors in the 13 differences are independent. Both implications can not be correct.

[2] Sasuly, M., *Trend Analysis of Statistics*, 1934, p. 181.

[3] Macaulay, F. R., *The Smoothing of Time Series*, 1931, pp. 46–50.

$$f_x = y_x + k_i$$

to this series, where

$$y_x = a_0 + a_1x + a_2x^2 + \cdots + a_rx^r$$

and k_i is one of p seasonal differences, repeating itself every pth term, so that

$$k_x = k_{x\pm p}.$$

Assume $n \geqq r+p$. The problem is to find the values of $a_0, a_1, a_2, \cdots, a_r$ and $k_1, k_2, k_3, \cdots, k_p$ which, taken simultaneously, minimize the function

$$z = \sum_{x=-l}^{l} (f_x - Y_x)^2,$$

with the condition imposed that $\sum_{i=1}^{p} k_i = 0$.

To develop certain general relationships, let us write

$$y_x = a_0 + u_x,$$

and

$$z = \sum_{x=-l}^{l} (a_0 + u_x + k_i - Y_x)^2 + \lambda \sum_{i=1}^{p} k_i \tag{1}$$

where λ is a Lagrange multiplier.[4] Then our solution must satisfy the equations obtained by setting the partial derivatives of z with respect to a_0 and each of the k's equal to zero. Setting the partial derivative of z with respect to some particular k_i equal to zero yields the equation

$$2m_i(a_0 + u_i + k_i - Y_i) + \lambda = 0 \tag{2}$$

where m_i is the number of times k_i occurs in the series $\sum_{x=-l}^{l} k_x$, and u_i and Y_i are the means of the values of u_x and Y_x, respectively, for those months or seasons in which k_i occurs. Summing equations like (2) for all the different k_i's yields

$$2\sum_{i=1}^{p} (m_ia_0 + m_iu_i + m_ik_i - m_iY_i) + p\lambda = 0.$$

This is equivalent to

$$2\sum_{x=-l}^{l} (a_0 + u_x + k_i - Y_x) + p\lambda = 0.$$

---

[4] See Osgood, W. F., *Advanced Calculus*, 1933, p. 180.

But setting $\partial z/\partial a_0$ equal to zero yields

$$2\sum_{x=-l}^{l}(a_0 + u_x + k_i - Y_x) = 0.$$

It follows that $\lambda = 0$. Hence, the terms in λ may be omitted from (1) and its derivatives, and the solution we seek may be obtained by partial differentiation[5] of

$$z = \sum_{x=-l}^{l}(a_0 + u_x + k_i - Y_x)^2.$$

By substituting zero for λ in (2) we find that

$$a_0 + u_i + k_i - Y_i = 0$$

or

$$k_i = Y_i - y_i \tag{3}$$

where y_i is defined as a_0+u_i; that is, y_i is the mean of the values of y_x for the months or seasons in which k_i occurs. This means that when the trend polynomial has been computed, the least squares solution for each seasonal difference can be obtained by finding the mean difference between the observed values and the computed trend for the months or seasons in which that particular seasonal difference occurs. By summing equations like (3) we find

$$\sum_{i=1}^{p} k_i = \sum_{i=1}^{p} Y_i - \sum_{i=1}^{p} y_i.$$

Since $\sum_{i=1}^{p} k_i$ is specified to be zero, this becomes

$$\sum^{p} y_i = \sum^{p} Y_i,$$

which is equivalent to

$$\sum_{x=-l}^{l}\frac{y_x}{m_x} = \sum_{x=-l}^{l}\frac{Y_x}{m_x}$$

where m_x is defined as the number of times the month or season corresponding to x occurs in the entire series of terms. Since

$$\sum_{x=-l}^{l}\frac{1}{m_x} = p$$

---

[5] The equation $\sum_{i=1}^{p} k_i = 0$ must be used in the solution, however, since the equations obtained from the partial derivatives are not independent.

in all cases, the least squares trend when r (the degree of the polynomial) is zero, is

$$a_0 = \frac{1}{p} \sum_{x=-l}^{l} \frac{Y_x}{m_x}. \tag{4}$$

To illustrate, take the simple problem of fitting a horizontal trend to a series covering 15 consecutive months beginning with May of one year and ending with July of the following year. In this case, $m_x = 2$ for May, June, and July of each year, since each of these three months occurs twice in the series. For each of the remaining nine months $m_x = 1$. Since $p = 12$, we have by substitution in (4),

$$a_0 = \frac{1}{12}\left\{ \sum_{-7}^{-5} Y_x/2 + \sum_{-4}^{4} Y_x + \sum_{5}^{7} Y_x/2 \right\}$$

or

$$a_0 = \frac{1}{24}\sum_{-7}^{-5} Y_x + \frac{1}{12}\sum_{-4}^{4} Y_x + \frac{1}{24}\sum_{5}^{7} Y_x.$$

To find the line of least squares for the same series, put

$$z = \frac{1}{2} \sum_{x=-l}^{l} (a_0 + a_1 x + k_i - Y_x)^2.$$

Setting the partial derivatives with respect to a_0, a_1, and each k_i equal to zero yields

$$\begin{array}{llr}
15a_0 & +k_1+k_2+k_3+k_4+2k_5+2k_6+2k_7+k_8+k_9+k_{10}+k_{11}+k_{12} & = \sum_{-7}^{7} Y_x \\
280a_1 & +k_1+2k_2+3k_3+4k_4-2k_5 \qquad +2k_7-4k_8-3k_9-2k_{10}-k_{11} & = \sum_{-7}^{7} xY_x \\
a_0-4a_1 & +k_8 & = Y_{-4} \\
a_0-3a_1 & +k_9 & = Y_{-3} \\
\cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \cdots \\
a_0+4a_1 & +k_4 & = Y_4 \\
2a_0-2a_1 & +2k_5 & = Y_{-7}+Y_5 \\
2a_0 & +2k_6 & = Y_{-6}+Y_6 \\
2a_0+2a_1 & +2k_7 & = Y_{-5}+Y_7
\end{array}$$

By subtracting $\sum_{i=1}^{12} k_i = 0$ plus one-half the sum of the last three equations from the first equation in this set, we obtain

$$12a_0 = \frac{1}{2}\sum_{-7}^{-5} Y_x + \sum_{-4}^{4} Y_x + \frac{1}{2}\sum_{5}^{7} Y_x.$$

Subtracting appropriate multiples of the last twelve equations from the second equation in the set so as to eliminate all the k's gives us

$$216a_1 = 6\sum_{5}^{7} Y_x - 6\sum_{-7}^{-5} Y_x.$$

It follows that

$$a_0 = \frac{1}{24}\sum_{-7}^{-5} Y_x + \frac{1}{12}\sum_{-4}^{4} Y_x + \frac{1}{24}\sum_{5}^{7} Y_x \tag{5}$$

and

$$a_1 = \frac{1}{36}\sum_{5}^{7} Y_x - \frac{1}{36}\sum_{-7}^{-5} Y_x. \tag{6}$$

Coefficients of polynomials of higher degree may be computed in a similar manner; but the procedure will frequently require considerable time and involve large numbers that are cumbersome to handle. On the other hand, when a general solution for the coefficients has once been obtained for some combination of r, p and n, and written down in a convenient place for future reference, the computation of a particular regression coefficient resolves itself into the simple procedure of applying a set of weights to the observed series.

Before deriving weights for fitting parabolas to seasonal data, let us consider briefly one or two devices that have been invented as aids in solving the problem of fitting polynomials to data that do not contain seasonal differences.[6] A very important contribution was the invention of the orthogonal polynomials. The first four of these polynomials may be written

$$G_{(0,x)} = 1,$$
$$G_{(1,x)} = x,$$
$$G_{(2,x)} = x^2 - \frac{n^2 - 1}{12}$$

and

$$G_{(3,x)} = x^3 - \frac{3n^2 - 7}{20}x.$$

In terms of these polynomials, the least squares regression curve may be written

---

[6] For a list of references to this problem, see Yule, G. U., and Kendall, M. G., *An Introduction to the Theory of Statistics*, 1937, references for Ch. 17, pp. 514–515.

$$A_0G_{(0,x)} + A_1G_{(1,x)} + A_2G_{(2,x)} + \cdots + A_rG_{(r,x)};$$

that is,

$$y_x = \sum_{i=0}^{r} A_iG_{(i,x)}. \tag{7}$$

These polynomials have the property that

$$\sum^{l} G_{(i,x)}G_{(j,x)} = 0$$

for all combinations of i and j except those where $i=j$. It follows that the least squares solution for the regression coefficient is

$$A_i = \frac{\sum_{-l}^{l} G_{(i,x)}Y_x}{\sum_{-l}^{l} G^2_{(i,x)}},$$

or

$$A_i = \sum_{-l}^{l} v_xY_x$$

where

$$v_x = \frac{G_{(i,x)}}{\sum_{-l}^{l} G^2_{(i,x)}}$$

Tables are now available[7] for the convenient computation of the coefficients.

By substituting equivalent values, (7) may be written in the form

$$y_x = \sum_{i=0}\left\{\sum_{x=-l}^{l} \frac{G_{(i,x)}Y_x}{\sum_{x=-l}^{l} G^2_{(i,x)}}\right\}G_{(i,x)}.$$

At the ordinate where $x=t$, this becomes

$$y_t = \sum_{i=0}\left\{\sum_{x=-l}^{l} \frac{G_{(i,x)}Y_x}{\sum_{x=-l}^{l} G^2_{(i,x)}}\right\}G_{(i,t)}.$$

---

[7] See Davis, H. T., *Tables of the Higher Mathematical Functions*, Vol. II, 1935, pp. 307–359; or Fisher, R. A., and Yates, F., *Statistical Tables for Biological, Agricultural and Medical Research*, 1938, table XXIII.

Collecting terms in each Y_x yields

$$y_t = \sum_{x=-l}^{l} w_{(r,x)} Y_x$$

where

$$w_{(r,x)} = \sum_{i=0}^{r} \frac{G_{(i,x)} G_{(i,t)}}{\sum_{x=-l}^{l} G^2_{(i,x)}} . \tag{8}$$

This formula is useful in deriving a set of weights which can be applied to the observed series of Y's to find the regression curve at any ordinate $x=t$. Sheppard's weights,[8] which are useful in fitting moving arcs to a nonseasonal series, can be obtained by setting $t=0$. Weights for computing the definite integral $\int_{t_1}^{t_2} y_x dx$ may be found by substituting $\int_{t_1}^{t_2} G_{(i,x)} dx$ for $G_{(i,t)}$ in (8).

Now weights identical with those obtained from (8) can also be derived directly by a procedure resting on the principle that, under certain conditions, the least squares solution is the unbiased linear estimate with minimum mean square error. Let ψ_t be a trend ordinate which it is desired to estimate. Assume that each observational error, $\epsilon_x = Y_x - \psi_x$, is a real number selected at random from a normal universe with mean zero and variance σ_ϵ^2. Let $y_t = \sum_{-l}^{l} w_x Y_x$ be an unbiased estimate of ψ_t (that is, let $\sum_{-l}^{l} w_x \psi_x = \psi_t$), and let $y_t - \psi_t = e_t$, the error in the estimate. Then the expected value of e_t^2 is $\sigma_\epsilon^2 \sum_{-l}^{l} w_x^2$. This is minimum when $\sum w_x^2$ is minimum. Hence, the unbiased linear estimate of ψ_t with minimum mean square error can be found by minimizing $\sum w_x^2$, with the necessary and sufficient conditions imposed to make $\sum w_x \psi_x = \psi_t$.

The proof[9] of the principle that the least squares solution is the unbiased linear estimate with minimum mean square error does not seem sufficiently general to cover cases where the fitted function includes seasonal differences as parameters—that is, cases where the equations obtained by partial differentiation of the sum of the squared residuals can not be solved by the method of determinants. It may be remarked, however, that as a criterion for solving most problems, minimum mean square error would appear to be at least as good as minimum mean square residual. The technique developed informally in the following paragraphs by using the mean square error criterion has led to solutions identical with the least squares solutions in the few cases where comparisons have been worked out.

---

[8] See Whittaker, E. T., and Robinson, G., *The Calculus of Observations*, 1924, pp. 291–299; Sasuly, M., *Trend Analysis of Statistics*, 1934, p. 46; Tintner, G., *The Variate Difference Method*, 1940, pp. 100–105.

[9] See Whittaker and Robinson, *op. cit.*, pp. 215–231.

First, compute a p-term moving total of the Y's. This results in a new series of observational values containing $n-p+1$ terms, $p-1$ terms being lost in the totaling procedure. This loss is offset by the elimination of p unknowns—namely, the p seasonal differences—which are equivalent to $p-1$ parameters because of the specification that their sum be zero. This new series contains all the parameters of the trend. Let one of the moving totals be represented by the expression

$$M_X = \sum_{x=X-h}^{X+h} Y_x$$

where $h=\frac{1}{2}(p-1)$. Let the estimated trend at the ordinate $x=t$ be written

$$y_t = \sum_{-L}^{L} W_X M_X \tag{9}$$

where W_X is a weight to be applied to M_X, and $L=\frac{1}{2}(N-1)$, N being the number of moving totals (that is, $N=n-p+1$). Then y_t is equivalent to $\sum_{-l}^{l} w_x Y_x$ where w_x is a linear function of the W's; for example,

$$\begin{aligned} w_{-l} &= W_{-L} \\ w_{-l+1} &= W_{-L} + W_{-L+1}. \end{aligned}$$

Now write

$$\begin{aligned} Z = \frac{1}{2}\sum_{-l}^{l} w_x^2 + \lambda_0\left(\sum_{-l}^{l} w_x - 1\right) + \lambda_1\left(\sum_{-l}^{l} w_x x - t\right) + \cdots \\ + \lambda_r\left(\sum_{-l}^{l} w_x x^r - t^r\right) \end{aligned}$$

in terms of the W's, and set the partial derivative of Z with respect to each W and each λ equal to zero. Solve these equations simultaneously to find the values of the W's, and substitute in (9).

Assume, for example, that we are to find a set of W's to apply to 15 terms in a monthly series to estimate the linear trend at the point $x=t$. In this case, $r=1$, $p=12$, and $n=15$. Taking twelve-month moving totals of the Y's results in four totals: $M_{-1\frac{1}{2}}$, $M_{-\frac{1}{2}}$, $M_{\frac{1}{2}}$ and $M_{1\frac{1}{2}}$. From the relationship

$$\sum_{-l}^{l} w_x Y_x = \sum_{-L}^{L} W_X M_X$$

we have

$$\begin{aligned}
w_{-7} &= W_{-1\frac{1}{2}}, \\
w_{-6} &= W_{-1\frac{1}{2}} + W_{-\frac{1}{2}}, \\
w_{-5} &= W_{-1\frac{1}{2}} + W_{-\frac{1}{2}} + W_{\frac{1}{2}}, \\
w_x &= W_{-1\frac{1}{2}} + W_{-\frac{1}{2}} + W_{\frac{1}{2}} + W_{1\frac{1}{2}} \text{ for } -4 \leqq x \leqq 4, \\
w_5 &= \phantom{W_{-1\frac{1}{2}} + {}} W_{-\frac{1}{2}} + W_{\frac{1}{2}} + W_{1\frac{1}{2}}, \\
w_6 &= \phantom{W_{-1\frac{1}{2}} + W_{-\frac{1}{2}} + {}} W_{\frac{1}{2}} + W_{1\frac{1}{2}}, \\
w_7 &= \phantom{W_{-1\frac{1}{2}} + W_{-\frac{1}{2}} + W_{\frac{1}{2}} + {}} W_{1\frac{1}{2}}.
\end{aligned}$$

The function to be minimized is

$$Z = \frac{1}{2} \sum_{-7}^{7} w_x^2 + \lambda_0 \left(\sum_{-7}^{7} w_x - 1 \right) + \lambda_1 \left(\sum_{-7}^{7} w_x x - t \right).$$

Replacing each w with the equivalent expression in the W's and setting the partial derivative of Z with respect to each W and each λ equal to zero yields

$$\begin{aligned}
12W_{-1\frac{1}{2}} + 11W_{-\frac{1}{2}} + 10W_{\frac{1}{2}} + 9W_{1\frac{1}{2}} + 12\lambda_0 - 18\lambda_1 &= 0 \\
11W_{-1\frac{1}{2}} + 12W_{-\frac{1}{2}} + 11W_{\frac{1}{2}} + 10W_{1\frac{1}{2}} + 12\lambda_0 - 6\lambda_1 &= 0 \\
10W_{-1\frac{1}{2}} + 11W_{-\frac{1}{2}} + 12W_{\frac{1}{2}} + 11W_{1\frac{1}{2}} + 12\lambda_0 + 6\lambda_1 &= 0 \\
9W_{-1\frac{1}{2}} + 10W_{-\frac{1}{2}} + 11W_{\frac{1}{2}} + 12W_{1\frac{1}{2}} + 12\lambda_0 + 18\lambda_1 &= 0 \\
12W_{-1\frac{1}{2}} + 12W_{-\frac{1}{2}} + 12W_{\frac{1}{2}} + 12W_{1\frac{1}{2}} &= 1 \\
-18W_{-1\frac{1}{2}} - 6W_{-\frac{1}{2}} + 6W_{\frac{1}{2}} + 18W_{1\frac{1}{2}} &= t.
\end{aligned}$$

Eliminating λ_0 and λ_1 and then solving for the four W's results in the solution:

$$\begin{aligned}
W_{-1\frac{1}{2}} &= 1/24 - t/36. \\
W_{-\frac{1}{2}} &= 0. \\
W_{\frac{1}{2}} &= 0. \\
W_{1\frac{1}{2}} &= 1/24 + t/36.
\end{aligned}$$

Substituting these values of W in (9) yields an estimate of y_t identical with that obtained from (5) and (6). Variations in this technique that lead to the same result are possible.

As another example, consider the problem of fitting a third degree parabola to the same 15 terms as before. In this case, we have the four conditions:

$$\sum_{-7}^{7} w_x = 1,$$

$$\sum_{-7}^{7} w_x x = t,$$

$$\sum_{-7}^{7} w_x x^2 = t^2,$$

$$\sum_{-7}^{7} w_x x^3 = t^3.$$

Since there are but four moving totals, these conditions alone are sufficient to determine the weights. Substituting equivalent values of W_X in the four equations, we obtain

$$\begin{aligned}
12W_{-1\frac{1}{2}} + 12W_{-\frac{1}{2}} + 12W_{\frac{1}{2}} + 12W_{1\frac{1}{2}} &= 1,\\
-18W_{-1\frac{1}{2}} - 6W_{-\frac{1}{2}} + 6W_{\frac{1}{2}} + 18W_{1\frac{1}{2}} &= t,\\
170W_{-1\frac{1}{2}} + 146W_{-\frac{1}{2}} + 146W_{\frac{1}{2}} + 170W_{1\frac{1}{2}} &= t^2,\\
-648W_{-1\frac{1}{2}} - 216W_{-\frac{1}{2}} + 216W_{\frac{1}{2}} + 648W_{1\frac{1}{2}} &= t^3.
\end{aligned}$$

The solution is

$$\begin{aligned}
W_{-1\frac{1}{2}} &= -\frac{73}{288} + \frac{t}{2} + \frac{t^2}{48} - \frac{t^3}{72},\\
W_{-\frac{1}{2}} &= \frac{85}{288} - \frac{19t}{12} - \frac{t^2}{48} + \frac{t^3}{24},\\
W_{\frac{1}{2}} &= \frac{85}{288} + \frac{19t}{12} - \frac{t^2}{48} - \frac{t^3}{24},\\
W_{1\frac{1}{2}} &= -\frac{73}{288} - \frac{t}{2} + \frac{t^2}{48} + \frac{t^3}{72}.
\end{aligned}$$

If these weights are used to compute the y's and (3) is used to compute the k's, the residuals $y_x + k_i - Y_x$ in this example will all be zero. This is to be expected, since $n-(r+p)=0$.

The weights may also be written in the more general form

$$W_X = \frac{F_{(0,X)}}{24} + \frac{F_{(1,X)}}{36}t + \frac{F_{(2,X)}}{48}\left(t^2 - \frac{85}{6}\right) + \frac{F_{(3,X)}}{72}(t^3 - 38t) \qquad (10)$$

where the F's take the following values:

X	$F_{(0,X)}$	$F_{(1,X)}$	$F_{(2,X)}$	$F_{(3,X)}$
$-1\frac{1}{2}$	1	−1	1	−1
$-\frac{1}{2}$	0	0	−1	3
$\frac{1}{2}$	0	0	−1	−3
$1\frac{1}{2}$	1	1	1	1

By substituting in (9), we obtain

$$y_t = \frac{\sum F_{(0,X)}M_X}{24} + \frac{\sum F_{(1,X)}M_X}{36}t + \frac{\sum F_{(2,X)}M_X}{48}\left(t^2 - \frac{85}{6}\right) + \frac{\sum F_{(3,X)}M_X}{72}(t^3 - 38t).$$

Hence,

$$y_x = A_0H_0/D_0 + A_1H_1/D_1 + A_2H_2/D_2 + A_3H_3/D_3 \qquad (11)$$

where

$$\begin{array}{lll} A_0 = \sum F_{(0,X)}M_X, & H_0 = 1, & D_0 = 24, \\ A_1 = \sum F_{(1,X)}M_X, & H_1 = x, & D_1 = 36, \\ A_2 = \sum F_{(2,X)}M_X, & H_2 = x^2 - 85/6, & D_2 = 48, \\ A_3 = \sum F_{(3,X)}M_X, & H_3 = x^3 - 38x, & D_3 = 72. \end{array}$$

Expressing y_x in the form shown in (11) is convenient because the terms are independent; that is, the formula for the least squares polynomial of r degree is

$$y_x = \sum_{i=0}^{r} A_i H_i/D_i,$$

A_i, H_i, and D_i being independent of r. Hence, after a polynomial of r degree has been fitted to the data, a polynomial of $r+1$ degree may easily be fitted without recomputing the coefficients already found. In this respect, (11) is like (7); but (7) is suitable only for fitting trends to nonseasonal data; or from another viewpoint, (7) is the formula for the special case when $p=1$. Note, also, that $F_{(2,X)}$ is not orthogonal to $F_{(0,X)}$, and $F_{(3,X)}$ is not orthogonal to $F_{(1,X)}$; that is $\sum F_{(0,X)}F_{(2,X)} \neq 0$, and $\sum F_{(1,X)}F_{(3,X)} \neq 0$.

The problem of fitting trends to seasonal data might be simplified if extensive tables of F, H, and D were computed and made generally available.

The minimum mean square error principle can be used to compute weights for graduating seasonal data with moving arcs. By substituting 0 for t in (10), or in each of the four equations preceding, we obtain

$$W_{-1\frac{1}{2}} = W_{1\frac{1}{2}} = -73/288,$$

and

$$W_{-\frac{1}{2}} = W_{\frac{1}{2}} = 85/288.$$

These weights may be used to find the midordinate of a second or third degree parabola fitted to 15 terms in a monthly series. The weights can also be obtained directly by minimizing

$$Z = \sum_{-l}^{l} w_x^2 + \lambda_0\left(\sum_{-l}^{l} w_x - 1\right) + \lambda_2\left(\sum_{-l}^{l} w_x x^2\right),$$

the odd-numbered λ's and all powers of t (above zero) dropping out of the function to be minimized. The solution is simplified somewhat by the symmetry of the weights.

Table I shows the weights for fitting moving arcs of the second or third degree to monthly data. The operators are surprisingly simple. The table includes a line to show the ratio of the standard error of y_0 to σ_ϵ when the ϵ's are random selections from a universe with mean zero. This is computed by finding the square root of $\sum w^2$. It affords some measure of the smoothing power of the weights, being comparable with $1/\sqrt{n}$ for the arithmetic mean, which is .289 for $n=12$.

When one has computed the W's to apply to the moving totals in estimating the trend at any particular ordinate, the equivalent w's can readily be found by substituting values. An alternative method is to take a p-term moving total of the W's with $p-1$ zeros at either end of the W-series. If a large number of interpolations is to be computed by the use of punch cards and Hollerith machines, it may be more convenient to apply these w's to the original data than to apply the W's to moving totals. In other situations, the computations will probably be simpler with the W's.

It should be understood that the purpose of this article is to suggest techniques and not to advocate their use without discrimination. If the seasonal variations are multiplicative, rather than additive, the techniques that have been described may not be appropriate. True, the multiplicative variations can be transformed into additive variations by simply taking logarithms of the original series; but in careful work, the procedure should not be governed by mere convenience. The best results are to be expected when the data are in such form, or have been put in such form, that they behave as though the deviations from the function whose parameters one is trying to estimate (not to be confused with the residual differences from the fitted function) were independent, homoscedastic, and normally distributed.

A METHOD FOR DETERMINING THE SIGNIFICANCE OF A SHORTAGE

BY LILA F. KNUDSEN*
Food and Drug Administration

THE METHOD given here for making quick decisions with regard to shortages in weight utilizes the range of a sample. There are occasions confronting the practical person where a choice has to be made between a rapid utilization of statistical methods or none at all. The method suggested here attempts to fill this need for simplicity and rapidity of application and can be used in any question of shortage—not necessarily weight only—but any suspected deviation from a standard or specification. This is a criterion for the average. There are other criteria to be considered—e.g. criteria dealing with single units in the sample—but these are not discussed here.

Suppose we are given a sample of several units from a lot having a definite declared weight; the actual weight of each unit is observed, whence the average weight and range of the sample can be easily computed. The question is "Shall we regard this lot as having the weight declared, or as having a definite shortage in weight?"

THE METHOD

Rule: Take the shortage in weight, and the range between the smallest and largest weights in the sample tested, and perform the simple division:

$$\frac{\text{shortage in weight}}{\text{range in sample}}.$$

Compare this ratio with the figures in the table corresponding to the number of units weighed (size of sample). (i) If the calculated value of this fraction is less than the value in Column A, the presumption is that no shortage exists. (ii) If this fraction lies between the entries shown in the two columns A and B of the table, shortage is suspected and additional units should be weighed. (iii) If the fraction is greater than the value in column B, the presumption is that a definite shortage exists.

---

* The author wishes to acknowledge gratefully the continued assistance and encouragement given by Dr. W. Edwards Deming of the Bureau of the Budget.

DERIVATION OF THE METHOD AND COMPUTATION OF TABLE I

Let

D = declared net weight
$\bar{x}$ = average net weight of sample
$\sigma_{\bar{x}}$ = estimated standard error of $\bar{x}$

Student's t is defined as

$$t = \frac{D - \bar{x}}{\sigma_{\bar{x}}}$$

where

$$\sigma_{\bar{x}} = \sigma/\sqrt{N}.$$

$\sigma_{\bar{x}}$ is the standard error of the average weight $\bar{x}$ of a sample of N units (estimated from the sample), and σ is the estimated standard error of single observations. The range of the sample can be introduced into the formula for t by noting that the average range in a sample of N is a multiple S of σ. The equation

$$\max - \min = \text{range} = S\sigma$$

will then be true on the average.[1] By replacing $\sigma_{\bar{x}}$ by (max − min)/S, the formula for t can be rewritten as

$$\frac{t}{S\sqrt{N}} \qquad \frac{D - \bar{x}}{\max - \min} \qquad \frac{\text{shortage in weight}}{\text{range in weight}}$$

Values of

$$S\sqrt{N}$$

are tabulated for various values of N in the table; they were calculated from the values of t that correspond to the two probability levels $P = .05$ in Column A, and $P = .001$ in Column B. These are approximations to the theoretical probability levels of the ratio (shortage in weight)/(range in weight).

A comparison of these tables with the A_2 values given on page 40 of the 1942 American War Standard (Ref. 2) shows that A_2 comes from a normal distribution while $t/S\sqrt{N}$ comes from the t distribution, the relation being

$$A_2 = (\text{normal deviate})/S\sqrt{N}.$$

---

[1] Values of S are available in Ref. 4, p. 165, or Ref. 8, p. 386, or as d_2 in Ref. 2.

TEST OF SIGNIFICANCE OF SHORTAGES

$$\text{Ratio} = \frac{\text{Shortage}}{\text{Range}}$$

N	A	B
3	1.468	10.777
4	.773	3.143
5	.534	1.655
6	.414	1.105
7	.342	.833
8	.294	.671
9	.259	.566
10	.232	.491
11	.212	.436
12	.195	.393
13	.181	.359
14	.169	.331
15	.160	.308
16	.151	.288
17	.143	.271
18	.137	.257
19	.131	.244
20	.125	.232
21	.120	.222
22	.116	.213
23	.112	.205
24	.108	.197
25	.105	.191
30	.091	.164
40	.074	.130
60	.056	.096
120	.035	.060
500	.015	.025

DISCUSSION

In any rule for coming to decisions, two types of mistakes may be encountered and are to be expected once in a while: either (i) a lot is accepted as passing a standard when further tests would not pass it; or (ii) a lot is rejected as not passing when further tests would pass it. The rule given here avoids these mistakes more successfully than any other workable scheme yet devised. Nevertheless, one can expect to make these mistakes once in a while.

As Simon (6) p. 115 quotes from Epictetus' "Discourses":

> Appearances to the mind are of four kinds. Things either are what they appear to be; or they neither are, nor appear to be; or they are, and do not appear to be; or they are not, and yet appear to be. Rightly to aim in all these cases is the wise man's task.

Here our "appearances to the mind" are the rules for coming to a decision, and Epictetus' last two kinds are the two types of mistakes we encounter. We are not worried about the instances where appearances and reality agree. A rule for coming to a decision attempts to make the appearance of things agree with reality more often than not. It is imperative to consider the consequences of the two types of mistakes and decide where to draw the borderline between them.

When the sampling is not random but merely representative, as is usually the case, there are no theoretical objections to the use of the range. However, even the theoretical discussions of the range of a random sample are on the whole very favorable toward its use (1) (4) (7) (8) and date back to the first part of the twentieth century. The range is very sensitive and its use in quality control work has had excellent success (2) (3) (5) (6).

APPLICATION

The following example illustrates the application of the method. A sample of 12 jars of peanut butter is taken from a lot. Each jar is labeled "8 oz. net weight." The individual net weights as taken are (in ounces):

8.2	7.7	7.5
8.0	7.5	7.4
7.6	7.3	7.5
7.6	7.4	8.0

The average of the 12 jars is 7.64 ounces, and the range in weight is $8.2-7.3=.9$. Applying the rule, we obtain

$$\frac{\text{shortage in weight}}{\text{range in sample}}=\frac{8.00-7.64}{.9}=.400.$$

This value is larger than the value .393 in the table for $N=12$, therefore a definite shortage in weight is presumed.

REFERENCES

(1) 1934—O. L. Davies, and Egon S. Pearson, *Supplement to the Journal of the Royal Statistical Society* 1, 76–89.

(2) 1941-2—(a) *Guide for Quality Control,*

(b) *Control Chart Method of Analyzing Data,*

(c) *Control Chart Method of Controlling Quality During Production.* Pamphlets (a) and (b) are bound together and appeared in 1941, 75¢; (c) appeared in 1942, 75¢. These standards are obtainable from the American Standards Association, 29 West 39 Street, New York.

(3) 1935—Egon S. Pearson, *The Application of Statistical Methods to Industrial Standardization and Quality Control* (British Standards Institution).

(4) 1931—Karl Pearson, *Tables for Statisticians and Biometricians* (Cambridge University Press, England).

(5) 1931—Walter A. Shewhart, *Economic Control of Quality of Manufactured Product* (Van Nostrand).

(6) 1941—Leslie E. Simon, *An Engineer's Manual of Statistical Methods* (John Wiley).

(7) 1927—Student, *Biometrika* 19, 151–164.

(8) 1925—L. H. C. Tippett, *Biometrika* 17, 364–387.

PAUL T. CHERINGTON, 1876–1943

PAUL T. CHERINGTON died in Philadelphia on the evening of April 24, 1943. That very morning in Washington he had participated in the deliberations of the Census Advisory Committee, a small committee of the American Statistical Association of which he had long been a member. The previous day had been spent by the same small group discussing the relationship of the Census Bureau to the statistical requirements of post-war planning. Overnight, Cherington evidently thought further about the problem. Immediately upon the opening of the meeting the following morning, he spoke for about fifteen minutes most sincerely and emphatically, voicing his profound concern with reference to the short-sightedness of business and government, the lip-service being given to post-war planning but the failure to reduce generalizations to the specific, and the essential requirements of statistical information on a national scale for adequate preparedness for the post-war period. It is a tragedy that these "last words" were not formally recorded as a living monument to his social vision. Characteristically, he was looking forward, working constructively to reduce broad social objectives to specific and practical terms.

Cherington always was a pioneer. He was among the first to recognize that marketing might be an area of recorded knowledge and his early teaching as a member of the original faculty of the Harvard Graduate School of Business Administration was to develop a systematic approach to marketing research. He later gave a practical demonstration of his thesis as Director of Research for the J. Walter Thompson Advertising Agency from 1922 to 1931. When the American Marketing Association was formed in 1931, he was the obvious choice for its first president. He founded the "Fortune" survey of public opinion, the pioneer enterprise in that presently well cultivated field. He was among the small group which fought for the so-called Census of Business in 1920, and spent endless hours in endeavors to increase its use and expand its usefulness. Additional items from his record would serve only to confirm the effectiveness of his creative pioneering. Nor can the complete record be found in his own activities. How many fires in other people were lit by sparks which he dropped, can never be known. But his friendliness, generosity and sincerity led to an enormous circle of friends and acquaintances who will always remember him with respect and affection.

In 1923, he drew up his specifications for a "successful life." And now it serves as a perfect summary of his own existence:

A successful life I should define as one which has brought a maximum of what have been called "durable satisfactions." Just what these are depends greatly on the individual. Measured by accepted social standards they would include a reasonable degree of personal comfort for one's self and his dependents, a sense of having lived helpfully with other people and a consciousness of having done something useful in the world which otherwise might not have been done. The surest aids to these accomplishments are an inquiring mind, clear thinking, an alert human sympathy, a lively sense of moral responsibility and a determination to work with those forces in the world which "make for righteousness."

WILLARD L. THORP

BOOK REVIEWS

Glenn E. McLaughlin
Review Editor

Studies in the National Income, 1924–1938, edited by A. L. Bowley. Cambridge: University Press. 1943. ix, 255 pp. $3.50.

This volume is the result of a study initiated in 1938 under the auspices of the National Institute of Economic and Social Research. Professor Bowley was Director of the investigation and, although others were associated with him on the project, he takes full responsibility for the methods followed and the opinions expressed. The outbreak of war forced curtailment of the original program and, as work cannot be resumed in the immediate future, the results obtained to date are being issued. While it is unfortunate that the intended plan could not be realized, the work as it stands is a major contribution to national income literature.

The volume consists of four chapters: The Definition of National Income; Some Constituents of the National Income; The Census of Production; Price Movements: Index of Real Income.

The chapter on definition by Dr. L. J. Elsas is a survey of the definitions used by the principal authors in the field, including the official estimates of those countries which issue national income data. The procedure used is to set up a comprehensive list of questions covering alternative concepts for the components of national income, and then to state the answers to the questions given in the works of the authorities listed. Thus, in highly condensed form one can see the degree of agreement among the twenty-four leading authors (including official estimates). It might be added that the questions answered for the principal authors are only those explicitly contained in their published works. An additional step might have been to indicate the answers to some of the questions implicit in the sources, contents, and procedures used by the authors, even though the questions were not formally discussed by them.

In the second chapter, Professor Bowley's paper from the *Journal of the Royal Statistical Society*, 1940, Part IV, is reprinted with some minor changes. This paper contains his estimates of national income for the period 1924–38. In general, it appears that Professor Bowley has adhered to the procedures he used with Lord Stamp in their estimate for 1924; his new results do not differ significantly from those of the earlier study for that particular year. A most useful addition to the *Journal* article made in this volume is a reconciliation of his national income aggregate for 1938 with that presented in the White Papers on war finance. As a consequence the projection of the White Paper estimates backward to 1924 by means of Professor Bowley's figures for the intervening years can be done much more confidently.

Further in the chapter Professor Bowley attempts a reconciliation of the labor force statistics of the Census of Population with the unemployment and insurance records. It will be recalled that it was largely through reliance upon the latter that Mr. Colin Clark arrived at a national income figure for 1924 differing from that of Bowley-Stamp (apart from differences of definition). The apparent conclusion of Professor Bowley here is that the unemployment insurance records present no evidence strong enough to induce him to alter his estimate derived from the Census of Population data. His analysis seems to indicate, in any event, that the differences between the two sources are not great. The reviewer must confess that his knowledge of the sources is far from intimate enough to warrant any judgment on the matter.

In the third chapter, Professor Bowley discusses the estimation of national income from production data. He starts with the estimate of Mr. Flux for the year 1924 and shows that, after comparability has been achieved, the result reconciles favorably with his own estimate derived by the summation of income method. Throughout the second and third chapters the technical excellence which one associates with Professor Bowley's name is everywhere in evidence. The reader is not only presented with results but is led through the appraisal of alternative sources and methods, and is shown the evidence for the leading assumptions adopted. One is always conscious of the meticulous attention to detail and of the balanced judgment that lies behind the final result. The one serious limitation of the study is that it is directed toward establishing the correct total for national income to the neglect of the breakdowns required for economic analysis. Of course, this may be due to the interruption of the work by the war.

The index of real income presented in Chapter IV constitutes the most novel departure in the volume. In addition to one derived from income totals deflated by a price index, as is customary, Professor Bowley has attempted a quantity index from various quantity series directly. While this method must seem like the green fields that lie beyond to anyone who has struggled with inadequate price data, it soon becomes apparent that it is no less question-begging than the customary procedure. In the absence of quantity data for much of services output, for example, Professor Bowley uses employment to measure changes over time, assuming that output per head of employed persons remained unchanged for the period 1924–38. This is equivalent to saying that the increase in efficiency was equal to the decrease in hours worked. While the reviewer cannot speak of the plausibility of that assumption for Great Britain, the result of the method would considerably understate the increase of services output in the United States. However, the difficulties involved are more deep-rooted. For insofar as services are intermediate products (transportation, retail trade, etc.) it is difficult to see how the change in number of persons employed has any relevance whatsoever. Yet, when quantity indexes for goods and for services are combined by means of "value added" weights in the base year, such intermediate products must influence the aggregate quantity index which is designed to measure changes solely in the output of final products. If the economy has, let us say, learned to economize on raw materials transportation, that is no reason why a given increase in the quantity of commodities reaching consumers should, in effect, be adjusted downward when the change in the latter is the object of measurement. Once it is recognized that only final product services can be included in the services indexes, the entire method loses much of its attractiveness.

The book contains much other interesting material which has not been touched upon in this brief review.

MILTON GILBERT

Bureau of Foreign and Domestic Commerce

Income Size Distributions in the United States, Part I. New York: National Bureau of Economic Research, by the Conference on Research in Income and Wealth. 1943. xxvi, 131 pp. $1.00.

In 1938 the Conference on Research on Income and Wealth appointed a committee, under the chairmanship of Milton Friedman, to prepare a critical summary of recent significant studies of personal income distribution. The proposed study was to be an analysis of the limitations of data and methodology and was expected to serve as a basis for coordinating future studies. The present volume constitutes the committee's summary of findings and recommendations. Anyone engaged in appraising the results of past studies or in planning future studies will find this report indispensable.

The critical review is introduced by a chapter on "The Why and How of Distributions of Income by Size" by Simon Kuznets. This chapter discusses in general terms the reasons for preparing income distributions: the relationship of size of the recipient's income to his spending of it, the bearing of size of income upon the recipient's activities and attitudes, and the factors determining income size distribution. The discussion, while concerned primarily with demonstrating the necessity of varying the concepts of income, recipient unit, and time period in accordance with the specific purpose of each investigation, also stresses the importance of planning each inquiry to achieve comparability among studies. Only in this way can there be a full exploitation of limited data.

The committee illustrates the general remarks of the introductory chapter by analyses of 16 important studies. The summaries in Part I have been adapted from very detailed studies which appear in Part II, available only in mimeographed form. A convenient tabular summary sets forth for each study the period covered, geographic coverage, population groups to which the income data apply, and the nature of the income data. A similar summary is provided for the definitions of the family unit, types of income, and types of receipts.

An analysis of the major shortcomings of these studies provides the basis of recommendations for cooperative action in future studies. Without attempting to force income data collected for a variety of purposes into a straight-jacket, the report urges a minimum degree of standardization and presents a reasonable outline for achieving it. The core of its recommendation is that an income concept for a particular purpose be devised and the tabulations planned so that comparisons can be made with a "standard" concept which is defined by the committee in considerable detail. In this manner a bridge would be erected linking future studies.

DANIEL CREAMER

Bureau of Foreign and Domestic Commerce

The Impact of Federal Taxes, by Roswell Magill. New York: Columbia University Press. 1943. ix, 218 pp. $3.00.

Professor Magill's lively volume makes its major contribution in three informative chapters on the impact of the tax law upon various forms of property disposition. Also, there is a useful sketch of the judicial history of Federal taxation in the thirties and a suggestive survey of the workings of the Bureau of Internal Revenue. The high competency of the author is evident as long as the discussion moves on the legal and administrative plane. Unfortunately, however, the more technical chapters are preceded by a general appraisal of the Federal tax system during and after the war, the level of which falls far below the standard set in the remainder of the book.

The chapters on the personal income tax, the estate tax, and the corporate income tax present an excellent survey of the impact of the statutes and their judicial interpretation upon various types of property settlements. Extensive reference to court decisions is made and will prove helpful to readers unfamiliar with the legal phases of taxation. The author's proposals for changes in the statutes are generally sound and would make for more equitable taxation. Thus, Professor Magill favors the taxing of Government securities and elimination of the advantages enjoyed by taxpayers in community property states under the Federal Income Tax. He emphasizes the tax advantage obtained by splitting the ownership of income earning assets. While there is no definite conclusion with respect to mandatory joint returns, the indications are that the author would be in their favor. The lack of correlation between the Federal estate and gift taxes is justly criticized. In their place a Federal tax on transfer is proposed which would cover inter vivos as well as testamentary transfers, with rates depending upon the amount of property received and the relationship of the recipient to the transferor. Accepting the conduit theory of corporate taxation, Professor Magill suggests that double taxation of dividends be eliminated by giving a credit to the corporation for personal income tax payment on dividends or vice versa. The importance of the tax structure with respect to corporate organization is emphasized in the conclusion that "the character of corporate structure has been more largely determined for some years by the tax provisions than by the industry of the anti-trust division of the Department of Justice."

In his technical chapters Professor Magill demonstrates the importance of the statutory and judicial phases of taxation for a realistic understanding of tax policy. The reader trained in economics will be startled to find, however, how little understanding of the broader economic aspects of taxation is evidenced in the first chapter of the book. The discussion in this chapter suffers from a complete failure to distinguish between the economic setting of taxation during the war and in the postwar economy. Although individual taxes are considered largely with reference to the war economy, Professor Magill on the very same basis somehow arrives at a proposed Federal tax budget for the postwar period. Of the total yield of $20 billion, $6.6 billion are to be derived from sales taxes including a general sales tax of 5 per cent and $8 billion are to be derived from the individual income tax, featuring a normal rate of 25 per cent and substantially reduced exemptions! Whatever the merits of this tax system for the war budget, it could hardly be acceptable for the postwar economy.

Economists may at times be guilty of disregarding some of the finer legal or administrative points of tax policy. But it is equally evident that tax policy in its broad outlines must first of all be based on an understanding of the *economic* issues involved. To this reviewer, the chasm between the legal-administrative and the economic training in tax problems appears most unfortunate.

Richard A. Musgrave

Washington, D. C.

Tax Yields: 1941. Philadelphia: Tax Institute, University of Pennsylvania. 1942. xi, 171 pp. $2.50.

This is the fifth volume of an annual series. The data include a survey of Federal tax collections for each of the years from 1913 to 1942. The annual tax collections of the states are presented for the years from 1938 to 1941. A monthly breakdown of state collection data is available for the first five months of 1941 and 1942.

Such a compilation is essential and necessary spadework. It is an important piece of groundwork which can serve as a starting point for a number of research projects. It would have been interesting to see a comparison of collections and expenditures. Revenue data alone are economically not very meaningful *per se*, although they are useful supplementary information. As the author points out, the tax collection figures alone do not give a comprehensive impression of the relative functional importance of state and local governments. Federal and state grants which are not listed in this volume, have increased during recent years. While the ratio of tax collections to national income is given for a period of thirty years, the reader misses a presentation of the revenue and expenditures ratio.

The data for the first five months of 1941 and 1942 can serve as clues for later war developments in tax collections in a limited way only. On one hand, rationing has been felt on a larger scale in more recent months, thus influencing gasoline tax, sales tax, and various excise tax collections. On the other hand, the upward trend of prices continued, partly mitigating the reduction in collections due to rationing, at least as far as sales taxes are concerned.

On the whole, the book is most useful. It is reliable, well organized, and comprehensive within the limits set by the author. Moreover, no other book brings identical information.

The reviewer and many other economists appreciate the tedious and laborious task of Miss Mabel Walker who is responsible for these annual compilations.

HENRY SIMON BLOCH

University of Chicago

The Burden of British Taxation, by G. Findlay Shirras and L. Rostas. New York: The Macmillan Company. 1943. xiii, 240 pp. $3.00.

A statistical study of the burden of taxation is beset with many complications. Aside from the difficulty of obtaining exact data, there are various items for which estimates must be made and, if the real burden is to be envisaged, there is the elusive problem of the allocation of subsidies. Furthermore, what are the effects of governmental expenditures in general on taxpaying classes?

The purpose of the present study is to ascertain the money burden of taxation for the years 1937–38 and 1941–42 for a range of different incomes in Great Britain. The authors have brought forward with many additions and improvements the authoritative Report of the Colwyn Committee on National Debt and Taxation in 1927. Within the scope of the assumptions set and the limitations of data, the work is a model of statistical investigation. When it has been impossible to use exact figures, as for example, in computing the burden of indirect taxes, estimates have been made with painstaking care and discriminating judgment. The reader is fully warned with caveats in respect to the conclusions drawn from the analysis.

A taxpaying family of four (husband, wife and two children) with a single income, is selected as representative of the social family in England today, in contrast with a family of five in previous studies. The levels of income chosen range from £100 to £50,000. Approximately nine-tenths of all revenue of the Exchequer is included for 1937–38 and four-fifths for 1941–42, the main items of exclusion being business profits and rates on dwelling houses. In 1937–38 the distribution of the tax burden was not substantially different from that of the years following the first World War. Notable changes have come with the wartime program of taxation, particularly in the middle income groups. In prewar years, the distribution was markedly progressive in the upper range, e.g., £1000 and upwards, and was quite regressive on lower incomes of £300 or less. On middle incomes falling between these figures the burden was "light in comparison." The War burden is "notably" progressive on incomes above £250, and in the opinion of the authors is slightly more regressive on incomes below that amount. Direct taxes are now important on the middle income group and are very heavy for higher incomes. In the lower income range indirect taxes are much more significant. All in all the "increase in the taxation of income is the most striking development in British taxation and is nothing short of a revolution."

The method of computing the incidence of death duties in Chapter VI, was devised by Mr. N. Kaldor and is a departure from the traditional "insurance method." It is termed the "annuity" method and is used because of the impracticable nature of the insurance method for larger investment incomes. While criticism may be directed to the sweeping assumptions made by Mr. Kaldor, the results of his annuity method of computation appear plausible when viewed in the light of the extremely high income and surtax rates now in effect.

TIPTON R. SNAVELY

University of Virginia

Fiscal Planning for Total War, by William L. Crum, John F. Fennelly, and Lawrence H. Seltzer. New York: National Bureau of Economic Research. 1942. xxv, 358 pp. $3.00.

Too much emphasis cannot be placed upon the principles of war financing in view of the magnitude and complexity of the problems encountered in financing our participation in World War II. While much water has run over the dam since this volume was written, it remains a significant analysis of the problems of financing a costly war with a minimum of inflationary effects in the most equitable and effective manner.

The authors present their analysis with reference to the fiscal year 1943. They assumed a gross national production of $125 or $130 billion and a net national income of approximately $113 billion. To cover Federal spending of $75 billion, a tax program to raise $40 billion was suggested. It was thought that some such sum would be required to limit borrowing to noninflationary sources.

Three alternative tax programs are offered, the reader to take his choice of higher or lower corporate and personal income taxes, higher or lower excises, and a general sales tax if he wishes to place more emphasis upon consumption taxes. All of the programs call for $2 billion in new revenue from the social security taxes and $500 million additional from estate and gift taxes. Each of the programs would raise approximately $12 billion. To meet the goal of $40 billion, which the authors think desirable, another $10 billion would have to be found. Beyond hinting that even more severe tax rates would be required, the authors do not advance particular suggestions for new revenue except by implication as they appraise the various available taxes. It is their purpose to indicate the rough outlines of a tax program adequate to obviate inflationary borrowing and to leave to the reader the privilege of deciding upon the specific taxes he would employ.

The tax program is combined with suggestions to cut at least $1 billion from nonwar expenditures and to exploit forced loans beyond the limits of voluntary loans and wisely selected taxes. The authors are careful to present the disadvantages as well as the advantages of each revenue proposal. Here may be found many of the essentials of a tax reform program as well as the requirements of a war taxing program.

The Federal tax receipts in the fiscal year 1943 actually fell far short of the $40 billion recommended. Only $23 billion were obtained from the gross tax and miscellaneous receipts, as compared with expenditures of some $78 billion.

The volume presents the results of a careful study of the principles and problems of war financing and is well worth reading by every student of public finance.

ALFRED G. BUEHLER

University of Pennsylvania

Accounts Receivable Financing, by Raymond J. Saulnier, and Neil H. Jacoby. New York: National Bureau of Economic Research. 1943. xv, 157 pp. $2.00.

This is an invaluable handbook for commercial banks, commercial finance companies, and factoring companies on the supply side of the business financing market and to concerns engaged in manufacturing, wholesaling, servicing, retailing, constructing, and mining on the demand side.

In its broader aspects this treatise is concerned with the cross-sectional pattern of business credit demands and the adaptations that financing institutions providing short and medium term credit have made in response to changing demands for their services.

This particular study deals almost exclusively with accounts receivable financing, although an attempt has been made to show how it is related to other financing methods. It is not concerned with instalment notes receivable, but rather with the process of "cashing receivables" by purchasing the invoices as well as other schemes relating to its title. It lives up to the statement by the authors that it is concerned with "the development and present quantitative importance of accounts receivable financing; the nature of institutions extending credit on this basis; the main features of the services offered by them; the chief characteristics of the businesses obtaining financial support from these agencies; the methods followed in conducting such financing; and the charges levied, income earned, costs of operation, and profitability of the business."

The authors have done a scholarly job of collecting data, treating it statistically and interpreting the results with a minimum of verbiage. Statements of fact are supported with adequate and clearly classified tables and in the appendix is a description of the procedures followed in getting the information. Throughout the book a careful distinction is kept between "factoring," which involves notification (non-recourse) and "non-notification financing," which is concerned with the purchase or assignment on a recourse basis.

Contrary to most books, the authors have placed a summary of their findings at the beginning rather than at the end, so that the reader has a prevue of what to expect.

I have only one minor criticism and that concerns the use of the term "net worth." We accountants are trying to steer clear of those misleading words, although they still appear in financial statements, texts, and other sources.

While this study contains much of interest for others, it should certainly be on the desk of all bank officials whose institutions in the future will be concerned with unused lending capacity. In it they will find a relatively new avenue of credit extension. It is a growing field and holds great promise for the future as it is being applied to an ever-widening range of industries.

ROBERT D. AYARS

University of Pittsburgh

Employment in Manufacturing, 1899–1939, by Solomon Fabricant. New York: National Bureau of Economic Research, Inc., 1942. xix, 362 pp. $3.00.

The present volume, a companion piece to the author's earlier volume on the output of manufacturing industries, 1899 to 1937, completes his statistical analysis of manufacturers from the turn of the present century to the inception of World War II. The problems and methods of measurement as well as the primary measures themselves are presented in a series of technical appendices that comprise half of the volume. The material presented in the body of the book is more aptly described by the volume's sub-title, *An Analysis of the Relation of Employment to the Volume of Production.*

Despite the fullness of the technical appendices, nowhere is there a discussion of the production index most appropriate for measuring output per unit of labor. Perhaps the omission stems from a realization that the character of the production index actually makes little difference, particularly over long periods of time, in view of the limited choice imposed by the available data. The appendices however, do more than merely set forth the statistical data underlying the author's analysis. Henceforth, all investigators faced with the formidable problem of achieving a comparable series from the Censuses of Manufactures between 1899 and 1939 will be indebted to Mr. Fabricant for his systematic and comprehensive presentation of the host of minor discontinuities that characterize the Census data.

In the body of the book Mr. Fabricant marshals in a lucid manner an imposing array of evidence to show convincingly (1) that "the industries with particularly large increases in both employment and output between 1899 and 1937 tended to have exceptionally large declines in employment per unit, whereas those with under-average increases in both tended to reduce employment per unit less sharply than the average"; (2) that rapidly growing industries have characteristically "tended to effect particularly large reductions not only in labor per unit but also in wage costs per unit of product, in unit value added (that is, wage and overhead costs, including profits), and in selling price"—although why this has been characteristic of some industries and not others is not investigated; and (3) that "the largest increases in both jobs and output occurred in the major groups of manufacturing industries whose fixed assets, in terms of net book values, grew most rapidly [while] among the groups with relatively slight increases in output and employment are the industries with less-than-average increments to capital assets."

These are the major findings of statistical relationships. The chapter on the growth and decline of manufacturing industries largely confirms the earlier findings of Simon Kuznets and A. F. Burns on the retardation of growth. Such a summary, however, fails to do justice to Mr. Fabricant's careful workmanship and shrewd appraisal of the limitations of the underlying statistical materials and the deviations from the general pattern. He eschews the tempting path of easy generalization on the problems of technological unemployment. His discussion on this point emphasizes rather the inherent complexity of the interrelations of manufacturing with the other sectors of our economy. In this view an analysis of data on manufacturing alone cannot provide the answers to the problems posed by technological unemployment.

Compared with other investigations in the field the present study may be said, to its advantage, to steer a middle course. The work of the National Research Project,[1] for example, was given over entirely to the construction of measures designed to answer questions that are devoid of economic content. For example, how much employment would there be if we again attained the production levels of 1929? At the other extreme is the study of The Brookings Institution[2] which engages in an oversimplified analysis of the distribution of the benefits of productivity gains among the factors of production. The conclusions have little or no basis in the statistical materials. Mr. Fabricant, on the other hand, illuminates much more of the problem by concerning himself with relevant and meaningful relationships yet restricting himself to conclusions that can be demonstrated.

DANIEL CREAMER

Bureau of Foreign and Domestic Commerce

---

[1] *Production, Employment and Productivity in 59 Manufacturing Industries, 1919–36*, by Harry Magdoff, Irving H. Siegel, and Milton B. Davis, WPA National Research Project (May 1939).

[2] *Productivity, Wages, and National Income*, by Spurgeon Bell, The Brookings Institution (Washington, D. C. 1940).

The Measurement of Business Activity in the Puget Sound Area, by Grant I. Butterbaugh. Chicago: Studies in Business Administration, School of Business, University of Chicago. 1943. v, 72 pp. $1.00.

In his study, *The Measurement of Business Activity in the Puget Sound Area*, Professor Butterbaugh has presented a local index which shows monthly changes in business in an important industrial and trading region of the country. Dominated by the lumber industry and leaning heavily on foreign trade, the Puget Sound Area differs widely from other industrial sections in the United States.

As expected, the author was confronted with a scarcity of regional data. Moreover, most of the series which were available extended back only a relatively few years which reduces somewhat the confidence in the adjustments for seasonal changes and for trend. To throw more light on business changes in the region, however, a second index was constructed without adjustment for trend.

Some less desirable series had to be used in the index because more adequate data were not available. For example, factory employment was included in the production series as a partial measurement of production. Clearly, man-hours is a better indicator of production; but such data are not available for the area. Likewise, department store sales were used to reflect retail trade, whereas the addition of other types of retail stores would greatly enrich the series. But here again, data are lacking.

Perhaps the author goes a little too far in refinement, when he assigns weights in fractional points, such as 5.25 points to gasoline sales and 8.75 points to department store sales. And the assumption that gasoline sales in the Puget Sound Area follow the same pattern as the total for the state may or may not be valid, since the assumption disregards the factor of distance in the less densely populated sections.

The composite index should reflect changes in business in the Puget Sound Area, owing to the fact that it embraces production, shipping, trade, and finance. Extreme care was used in the selection of series in order to eliminate so far as possible duplications and at the same time to include sufficient data to reflect local conditions. It is true, of course, that business in any one region of the country follows a pattern more or less like that in the United States as a whole; but there are significant differences. Regional indexes are exceedingly helpful in understanding the over-all picture of business in the country, and the Puget Sound index is a welcome addition to the few now available.

BERVARD NICHOLS

University of Pittsburgh

American Agriculture, 1899–1939: A Study of Output, Employment and Productivity, by Harold Barger and Hans H. Landsberg. New York: National Bureau of Economic Research. 1942. xxii, 440 pp. $3.00.

This book presents an intensive study of the annual volume of net agricultural output in the United States, by individual items and overall, of labor used in agriculture, of food consumption, and of farm productivity per acre and per man.

While the basic data are valuable, the analysis suffers from a number of weaknesses:

(1) The limitation to the period since 1899, and the general failure to take into account agricultural history prior to the 1900's.

(2) The use of linear trends without reference to whether they were adequate to represent the data.

(3) Lack of familiarity with public agricultural policy during the 1930's. With one minor exception, public policy is treated as a matter of A.A.A. restrictions, without recognition of the soil conservation emphasis since 1936, or its possible relation to subsequent increases in yields and output.

(4) Elaborate calculations of labor requirements for crop and livestock production (pp. 254–277) carried far beyond the significance of the basic data.

In the final chapter, the authors develop the thesis that the rapid secular increase in productivity per capita, both in agriculture and in industry, together with the relatively inelastic demand for food, forces a continuous reduction in the proportion of workers engaged in agriculture, and a resultant chronic depression in the relative economic position of farmers. C. Reinold Noyes, Director of the National Bureau, vigorously challenges these conclusions and their significance as a forecast for the future in an extensive dissent on pp. 316–321. He bases his dissent on the imperfections of analysis such as those noted above and especially on the failure to take into account the historical shift from pioneer settlement and expansion to a settled country. While the reviewer agrees with Noyes as to the weaknesses of the analysis, he still is inclined to agree with the general drift of the authors' conclusions, having reached parallel conclusions several years earlier on the basis of a superficial study covering a longer historical period.[1]

The book will be useful to research workers for its copious reference tables of the annual net output, prices, and consumption of individual farm products, for its new index numbers of net physical production, which take into account changes in farm livestock inventories, and for its careful canvass of the comparability of the available data on agricultural employment.

MORDECAI EZEKIEL

U. S. Department of Agriculture

---

[1] Mordecai Ezekiel, "Population and Unemployment," *The Annals of the American Academy of Political and Social Science*, Vol. 188, pp. 230–242. Nov., 1936.

Western Prices before 1861, A Study of the Cincinnati Market, by Thomas Senior Berry. Cambridge: Harvard University Press, Harvard Economic Studies, 74. 1943. xxi, 645 pp. $5.00.

This excellent study sheds much light, not only upon the behavior of prices in the West and particularly in Cincinnati, but also upon the general pattern of economic life in Cincinnati and its tributary area. Attention is called to the long-term trends in the growth of the West and in the declining cost of manufactures and imported goods as compared with the values of exportable agricultural produce of the area. Improved transportation and better production techniques are emphasized. The behavior of markets and prices during the period when dependence was placed upon river transport is shown to differ from the more marked seasonal pattern prevailing at a later date when the railroad had become well developed. Moreover, stress is laid upon the influence of traffic development in determining the character of transport facilities.

Discussion of the salt trade leads to an interesting detailed account of the efforts of the saltmakers to limit production and control prices. The close relationship between the Kanawha salt companies and the later Michigan salt associations is pointed out as is the later borrowing of technical methods by the oil industry and the probability that technical methods were also transferred. This section lends more support to the theory that the combination movement after 1880 was but an extension of earlier and not unsuccessful attempts of similar nature.

In his survey of the cyclical behavior of prices, the author sheds much new light upon banking and currency in the West. Moreover, through a detailed examination of successive crises in a particular area he strengthens our grasp of the nature of these phenomena during the period under discussion. It is possible to view them less abstractly, not in terms of averages, but in the light of their impact upon the life of a particular area. Much attention is given to the note issue of the various financial institutions serving the area, the difficulties attending the use of paper currency, the stringency of specie, and the periodical tightness in the supply of New York exchange. The author advances for further investigation the interesting hypothesis that the relative value of any single issue of bank notes depended more upon the total amount of coin held in the issuing bank than upon the quantity of the notes in circulation.

The mass of detailed information drawn together from widely scattered sources is well organized and ably presented. The study is capped by a general summary and conclusions which bring together the major points in the analysis. It merits the attention of all economists concerned with the general range of problems treated.

ERNEST W. WILLIAMS, JR.

War Production Board

International Agreements on Conservation of Marine Resources, by Jozo Tomasevich. Stanford University: Food Research Institute, 1943. xi, 297 pp. $3.00.

The author conducts a critical examination of the problems involved in international conservation of marine resources and then follows with the history and development of international agreements to conserve and sustain the fisheries for three biologically different species, namely, fur seals, halibut, and sockeye salmon—all found in the North Pacific. At present the fisheries for the first two species are under international control and regulation while the third is undergoing intensive scientific study preparatory to regulation. The commodities resulting from the three fisheries differ greatly. The fur-seal herd supplies a valuable type of fur; the halibut fishery, a food fish sold exclusively in fresh or frozen state; and the sockeye-salmon fishery supplies a product marketed almost entirely in the canned form.

The underlying problems are grouped under three headings: (1) international law, (2) scientific research, and (3) conservation and systematic development and use of fisheries as such.

One of the defects requiring correction in the existing international agreements on fur seals, halibut, and sockeye salmon, as pointed out by the author, is that they still leave these fisheries open for invasion by "foreign" fishermen, that is by fishermen native to countries not a party to the agreement. These agreements seem to have recognized the old Roman law which regarded the sea as common domain, free for the use of all and its riches were *res nullius*—unappropriated goods. Property rights in the products of the seas could be established and kept only with actual possession.

The author believes that it will likely be necessary in the future to give more consideration to the economic and technological problems involved in the development of international fishery agreements. Thus far emphasis has been placed on the biological side. Many will agree with his belief since fishery enterprises are held in delicate economic balance and any controls which regulate the intensity, amount, time, and place of fishing may necessitate economic adjustments involving radical changes in fishing gear and vessels and involving an altered pattern of labor supply and prices. These aspects are analyzed and an evaluation is made of the effect of the different types of controls. The author also points out the difficulties experienced by each industry, especially the halibut industry, in adjusting its activities to conform with the controls.

Under the fur-seal agreement the fur-seal industry became a U. S. Government monopoly with the high contracting parties sharing in the proceeds as agreed in the convention. Under the halibut agreement the industry continued as a multitude of small enterprises, but largely limited to that group of fishermen which engaged in the fishery prior to its regulation and control by the halibut commission. As the reader considers the nature, aims, techniques, limitations, and achievements of the international agreements on fisheries as set forth in the book, the question may arise in his mind as to whether these agreements—which are aimed at conserving, sustaining, and increasing supply for the public good—are of primary benefit to the interests engaged in the fishery enterprise, consumers of the product, or both.

There is no doubt but that government officials charged with the responsibility of drawing international fishery agreements, as well as fishery industrialists and students will find a wealth of interesting and useful material for reference and study in Dr. Tomasevich's well-written book.

REGINALD H. FIEDLER

Office of Foreign Economic Administration

Public Trusteeship, by Norman S. Heaney. Baltimore: John Hopkins Press. Studies in Historical and Political Science, Series LX, No. 4. 1942. 130 pp. $1.50.

This monograph explores in a very competent fashion a field of public enterprise so little known in the United States as probably to require some explanation. By public trusteeship Heaney refers to the performance by a government official of the trust function for private purposes as defined by English common law. Public trusteeship is thus typically a phenomenon of England and her colonies and dominions. Heaney's study is primarily an examination of the working of the office of the English Public Trustee, a government department which has been in continuous operation since 1908. One chapter, however, is devoted to experience with a similar institution in New Zealand, dating from 1872, which offers an interesting contrast to the English trustee in that a number of other fiduciary duties (serving as executor, administrator, guardian, receiver, liquidator, etc.) have been added to those of trusteeship.

So far as the English Public Trustee is concerned, Heaney traces the acceptance of the idea through a long period of consideration in Parliament, and then turns to a detailed analysis of the types and volume of business handled and an examination of practices—organizational, administrative, and investment—in the management of trusts. While there is no deliberate attempt to place the work of this agency against the larger background of British experience with other public enterprises, the information necessary for such a comparison is fully given. It appears that the original policy of the Trustee was a vigorous promotional one, which was assisted by an absence of serious competition from private corporate trustees. However, the office has more recently adopted a passive attitude, and Heaney believes that it is being choked off by private competitors. The staff of the Trustee was originally outside the Civil Service, but was later brought in over the objection of the Trustee, who feared the effect of this step on the competitive spirit of the agency. Fees charged are expected to cover the cost of the office, and have done so continuously since 1923. Comparison with the costs of private trustees is impossible, owing to absence of data on their operations.

In the concluding chapter of his thorough and scholarly study, Heaney examines the problems that would face a system of public trusteeship in the United States, which he regards as a "clear path to trust improvement" though not a "crying need."

C. HERMAN PRITCHETT

The University of Chicago

The Tragedy of European Labor, 1918–1939, by Adolf Sturmthal. New York: Columbia University Press. 1943. xii, 389 pp. $3.75.

As indicated by the title, Dr. Sturmthal has here presented a picture of the failures of organized labor in Europe between the two wars. That picture is excellent; calmly drawn, judiciously weighed, based upon reliable facts, and generally valuable. Students of history, political events, or of the labor movement, as well as labor leaders themselves can well afford to put time and thought into this comprehensive study.

The key to the whole study is the distinction between "pressure-group action and political activity"; and it is upon this basis that the author makes his penetrating examination of labor's history as a cultural phenomenon during more than twenty fateful years.

The first section of the study is devoted to giving the explanation for the failure of labor in terms of the criteria established by the author. That failure is seen as being a failure by European democracy to integrate social groups with their conflicting interests into a functioning industrial society. At this point labor, though strongly committed to the defense of democracy did not, or could not, stem the progressive disintegration of democracy. Thus the stage was set for the Fascist reintegration in terms of force.

Labor inherited the reins of government in many places after the first World War, but for several clearly illustrated reasons the various labor governments were unable to grow, as it were, with the needs of the era, and to provide the necessary political leadership. Labor, as a political party, was a pressure group because it did not make an effort to integrate its interests with those of the community of which it formed a part. Democracy could not prosper without the "active participation of labor," while labor, even when in power, was refusing its real responsibilities. The result, democracy fell, and with it went labor.

European labor, as a pressure group, was concerned only with a "narrow range of problems, namely those immediately affecting the interests of its membership." Thus, when in power, labor governments confined themselves to trades-union questions such as wages and hours, and did not bother to enter and influence the larger national and international issues upon which the fate of the continent rested.

Moreover, labor governments often relied upon traditionally conservative elements of pre-war society in utilizing the old civil service and the old Imperial army corps officers as props for the new democratic states. This was particularly the case in the young German democracy when Rightist Free Corpsmen defended the Socialist government against striking workers.

A tragically "amusing" factor in the failure of labor at this time, also, was the spectacle of labor governments lending naive support to *laissez faire* doctrines because of their theoretical view that the dialectic of the inevitable downfall of capitalism demanded that no attempts be made to work within the existing economy to affect any peaceful changes. Thus labor viewed plans for shortening the depression with alarm on the grounds that depressions were an inevitable part of capitalistic economy and that nothing should be done until capitalism fell.

In his concluding chapter Dr. Sturmthal observes that, "the progress of labor from the status of pressure group to political party could succeed only to the extent that the masses themselves gained understanding of their own tasks. . . . Only improved insight into the situations confronting them could induce the workers to abandon their narrow trades-unionist outlook in favor of broad constructive policies which would serve both the entire nation and create the conditions necessary for labor's further social advance." Unfortunately the record as here presented shows that labor did not gain this insight, although the British Labor Party, in forming the coalition government of 1940, seems to give signs of a new, constructive approach.

One of the most useful phases of this book is the attention paid to the significance of the integration of the various elements of modern industrial society. The author has clearly grasped the need for organization in terms of the new functional interdependence of men. The story of European labor is a most powerful plea for the necessity of integrated thinking and mutually constructive action for common goals which must transcend national boundaries if men in any one land are to be free from want and fear.

ROBERT I. CRANE

Washington, D. C.

Elementary Statistical Methods, by Helen M. Walker. New York: Henry Holt and Company, Inc. 1943. xxv, 368 pp. $2.75.

This is a text for a first-semester course in descriptive statistics for students who hope to conduct educational research or to read the literature on such research. The emphasis is on the development and understanding of ideas rather than on formulas and computational routines. At the same time the book is intended to serve as an introduction to the author's coming volume, *Statistical Inference*. The student is introduced to analysis of variance, the sampling distribution, and sampling problems as well as to averages, variability, frequency distribution, and simple correlation.

Well-chosen quotations and historical references which add to the interest of the book are frequently given. Mathematical notes and proofs of some of the formulas shown in the text are given in one of the five appendices where they may be more easily skipped by those not able to follow such discussions.

It is intended that the instructor will provide adequate drill material. The checking of problems worked out in table form in the book is suggested for practice. Moreover, a number of unusually good exercises are given throughout the book.

Beginning statistical workers who might expect this to be a reference book containing explanations of all the usual statistical techniques set forth in an elementary manner may be disappointed: the topics of partial and multiple correlation, measures of reliability, time series, and index numbers are not discussed. The list of formulas in the appendix are not especially adapted to ready reference since they are shown without names; nor are formulas for the standard deviation and the correlation coefficient shown in the forms best adapted to machine calculation.

A characteristic of the book which may be troublesome to beginning students is that statistical terms are used many pages before they are defined or explained.

The titles of the fifteen chapters, in order, are as follows: Orientation; The Nature of Measurement; The Frequency Distribution; The Presentation of Data in Tables; Measures of Relative Position; Symbolism in Statistics; Averages; Measures of Variability; The Mean and Standard Deviation under Various Circumstances; Further Characteristics of a Frequency Distribution; The Normal Distribution; Regression and Correlation—General Concepts; Regression and Correlation—Computations; Correlation and Regression—Interpretation; and Sample and Universe.

LAVERNE K. BURKE

Community Occupational Surveys, by Marguerite W. Zapoleon. Washington: U. S. Office of Education, Government Printing Office. 1942. vii, 199 pp. 25 cents.

This is a study of ninety-six surveys completed during the period 1930–40. Part I deals with such topics as the nature and history of surveys; then it takes up the preliminary research, planning the content, planning survey forms, community cooperation, training personnel, preparation of reports and results. Part II discusses in some detail twelve to fifteen of these surveys. The surveys that are covered in detail range from small and rural communities up to the occupational census of the United States for 1940. Many excellent comments are made on these particular surveys. Part III discusses the steps that should be taken in a community occupational survey. This section provides a very fine check list for anyone planning to study the occupations in a community. This outline of the steps taken should prove of great help in many situations.

This volume as a whole gives a very fine discussion of the survey techniques that are more commonly used. As this is all the volume claims to do, the following comments may not seem to be in order. On page 66 it is admitted that continuous surveys are really what we need. "What is needed, then, is a popular demand to set that machinery in motion and to supply adequate personnel and materials with which to carry on a continuous inventory of labor supply." It would seem as though the study of all of these surveys would have suggested many techniques for carrying on this continuous survey even with the resources now available. It is the opinion of the reviewer that any community that wants to do so, if it is willing to use the young people in the school, can immediately start a continuous survey program and keep it up to date at all times. This is probably the most realistic way that occupational information can be brought to the young people of a community.

It should also be the best possible kind of civic training showing them the importance of the working life of their own community. Almost all of these surveys contain special suggestions of shortcuts and methods that would be of help to other communities. It might have led to a great improvement in the technique of making surveys if these could have been brought together in one place.

In addition to the main body of the report, there is a very valuable appendix giving an annotated bibliography of reports of community occupational surveys. There is another appendix giving the bibliography of community occupational surveys and there is a third appendix giving samples of the actual cards and records used in making surveys. This volume should play an important part in bringing closer the day when we will have continuous and up to date information regarding the occupational supply and demand in every community in the land.

HAROLD F. CLARK

Columbia University

Mathematical Statistics, by S. S. Wilks. Princeton: Princeton University Press. 1943. xi, 284 pp. $3.75.

The study of Mathematical Statistics has been seriously hampered by the lack of good books and by the fact that many important developments of the theory can be found only in the original form and widely scattered in scientific periodicals. The author has performed a very valuable service to all interested in the field by writing the present excellent book. The content of the lithoprinted book is formed chiefly by the subject matter of a two-semester course given by the author for advanced undergraduate and beginning graduate students. As such it has, of course, its limitations. It was impossible to cover in great detail all important developments of the theory or to give very extensive illustrations. However, the book contains the basic material in a concise form and gives an excellent introduction to the more recent developments of the mathematical theory of statistics. At the end of the book a number of publications are listed which are recommended for supplementary reading. The book contains chapters on the following subjects: Distribution Functions, Some Special Distributions, Sampling Theory, Sampling from a Normal Population, On the Theory of Statistical Estimation, Tests of Statistical Hypotheses, Normal Regression Theory, Applications of Normal Regression Theory to Analysis of Variance Problems, On Combinatorial Statistical Theory, and An Introduction to Multivariate Statistical Analysis.

The reader will particularly welcome the chapters on Combinatorial Statistical Theory and on Multivariate Statistical Analysis. These theories are of growing importance and substantial advances in them have been achieved recently. Unfortunately, the literature on these topics is widely scattered and not sufficiently well known. Although the presentation of this material is necessarily brief, the reader will find it exceedingly useful. In the discussion of problems of analysis of variance (Chapter IX) the author brings out clearly the fact that these problems can be considered as problems in regression theory, a fact which is not so generally known.

Readers with some mathematical background will find the book an excellent introduction to the modern developments in the theory of mathematical statistics. Also teachers of mathematical statistics will find it very helpful in their classroom work. The book can be highly recommended to everyone interested in this field.

ABRAHAM WALD

Columbia University

PUBLICATIONS RECEIVED*

Backman, Jules. *The Price Control and Subsidy Program in Canada.* Washington: The Brookings Institution, Pamphlet No. 52. 1943. 68 pp. 50 cents.

Blakenship, Albert B. *Consumer and Opinion Research.* New York: Harper & Brothers. 1943. x, 238 pp. $3.00.

Bowley, A. L., Editor. *Studies in the National Income, 1924–1938.* New York: The Macmillan Company. 1943, ix, 255 pp. $3.50.

Brandt, Karl. *Fats and Oils in the War.* Stanford University: Food Research Institute, Stanford University, War-Peace Pamphlets No. 2. 1943. 57 pp. 25 cents.

Brody, Maurice S. *Wage Rates and Living Costs in a War Economy.* Chicago: The University of Chicago Press, Studies in Business Administration, Vol. XIII, No. 3. 1943. vi, 38 pp. $1.00.

Burnham, Elizabeth A. *Expenses and Profits of Limited Price Variety Chains in 1942.* Boston: Harvard University Graduate School of Business Administration, Bureau of Business Research, Bulletin No. 118. 1943. vi, 42 pp. $1.00.

Coats, R. H., and M. C. Maclean. *The American-Born in Canada.* Toronto: The Ryerson Press. 1943. xviii, 176 pp. $3.75.

Deardorff, Neva R., and Marta Fraenkel. Hospital Discharge Study, Vol. II, *Hospitalized Illness in New York City.* New York: Welfare Council of New York City. xvii, 349 pp. $1.00.

Gist, Noel P., C. T. Pihlblad, and Cecil L. Gregory. *Selective Factors in Migration and Occupation.* Columbia: University of Missouri, The University of Missouri Studies, Vol. XVIII, No. 2. 1943. 166 pp. $1.50.

Hurwitz, Howard L. *Theodore Roosevelt and Labor in New York State.* New York: Columbia University Press. 1943. 316 pp. $3.75.

International Association of Ice Cream Mfrs. *Ice Cream Sales Index, January 1st through December 31st, 1942.* Special Bulletin No. 67. Washington: International Association of Ice Cream Mfrs. 1943. 56 pp. $1.00.

Kells, Lyman M. *Calculus.* New York: Prentice-Hall, Inc. 1943. viii, 509 pp. $3.75.

Kleinhaus, H. I. *1942 Departmental Merchandising and Operating Results of Department Stores and Specialty Stores.* New York: National Retail Dry Goods Association. 1943. 55 pp. $5.00.

Koch, Albert R. *The Financing of Large Corporations, 1920–39.* New York: National Bureau of Economic Research. 1943. xiv, 141 pp. $1.50.

League of Nations. *Quantitative Trade Controls: Their Causes and Nature.* New York: Columbia University Press, Economic and Financial II. A. 5. 1943. 45 pp. 50 cents.

League of Nations. *Trade Relations between Free-Market and Controlled Economies.* New York: Columbia University Press, Economic and Financial II. A. 4. 1943. 92 pp. $1.00.

Lutz, Friedrich A. *International Monetary Mechanisms, The Keynes and White Proposals.* Princeton: Princeton University, Essays in International Finance, No. 1. 1943. 21pp.

McNair, Malcolm P. *Operating Results of Department and Specialty Stores in 1942.* Boston: Harvard University Graduate School of Business Administration, Bureau of Business Research, Bulletin No. 118. 1943. vi, 52 pp. $2.50.

Mekler, Ana. *El Costo de la Vida Ob-*

---

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. Some items of minor interest to statisticians have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications which are to be found in the *Revue de l'Institut International de Statistique, Journal of the Royal Statistical Society, American Economic Review, Population Index, Transactions of the Actuarial Society of America, The Record of the American Institute of Actuaries, Sankhyā., the Indian Journal of Statistics,* and *Estadística.*—Editor.

rera en America. Washington: Union Panamericana, Oficina de Informacion Obrera y Social. 1943. iii, 139 pp.

Myers, Charles A., and W. Rupert Maclaurin. *The Movement of Factory Workers.* New York: John Wiley & Sons, Inc. 1943. viii, 111 pp. $1.50.

National Industrial Conference Board, Inc. *Effects of Taxes Upon Corporate Policy.* New York: National Industrial Conference Board, Inc. 1943. 136 pp.

Ogburn, William Fielding. *War, Babies, and the Future.* New York: Public Affairs Committee, Inc., Public Affairs Pamphlets No. 83. 1943. 30 pp. 10 cents.

Richtmeyer, Cleon C., and Judson W. Foust. *Business Mathematics.* New York: McGraw-Hill Book Company, Inc. 1943. Second Edition. xv, 401 pp. $2.75.

Seidman, Joel. *Union Rights and Union Duties.* New York: Harcourt, Brace and Company. 1943. viii, 238 pp. $2.50.

Stetler, Henry G. *The Socialist Movement in Reading, Pennsylvania, 1896–1936, A Study in Social Change.* Storrs, Connecticut. 1943. vii, 198 pp.

Waugh, Albert E. *Elements of Statistical Method.* New York: McGraw-Hill Book Company. 1943. Second Edition. xxi, 532 pp. $4.00.

Whitaker, Arthur P. *Inter-American Affairs, 1942, An Annual Survey: No. 2.* New York: Columbia University Press. 1943. ix, 252 pp. $3.00.

Wickizer, V. D. *The World Coffee Economy with Special Reference to Control Schemes.* Stanford University: Food Research Institute, Stanford University. 1943, x, 258 pp. $3.00.

Woytinsky, W. S. *Earnings and Social Security in the United States.* Washington: Social Science Research Council, Committee on Social Security. 1943. xiii, 260 pp. $2.50.

Government Publications

National Resources Planning Board. *Demobilization and Readjustment.* Washington: Government Printing Office, Superintendent of Documents. 1943. 106 pp. 20 cents.

——— *Industrial Location and National Resources, December 1942.* Washington: Superintendent of Documents. 1943. vii, 360 pp. $1.50.

——— *Regional Planning, Part XI The Southeast.* Washington: Superintendent of Documents. 1943. viii, 250 pp. 55 cents.

State of Tennessee, Department of Public Health. *Annual Bulletin of Vital Statistics for the Year 1942.* Nashville: Department of Public Health. 1943. 80 pp.

United States Tariff Commission. *Foreign-Trade and Exchange Controls in Germany.* Washington: Government Printing Office, Report No. 150, Second Series. 1943. xi, 294 pp. 35 cents.

I.A.R.I. 75

INDIAN AGRICULTURAL RESEARCH
INSTITUTE LIBRARY, NEW DELHI.

Date of Issue	Date of Issue	Date of Issue

GIPNLK—H-40 I.A.R.I.—29-4- 5—15,000